Winter 3-3-2018

Cost-Effective Use of Follow-Up Colonoscopies and Fecal Immunochemical Tests in Colorectal Cancer Screening

Grace N. Joseph
Louisiana Tech University

Follow this and additional works at: https://digitalcommons.latech.edu/dissertations

Recommended Citation
https://digitalcommons.latech.edu/dissertations/30

This Dissertation is brought to you for free and open access by the Graduate School at Louisiana Tech Digital Commons. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of Louisiana Tech Digital Commons. For more information, please contact digitalcommons@latech.edu.
COST-EFFECTIVE USE OF FOLLOW-UP COLONOSCOPIES AND FECAL IMMUNOCHEMICAL TESTS IN COLORECTAL CANCER SCREENING

by

Grace Naomie Joseph, B.S.

A Dissertation Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

COLLEGE OF ENGINEERING AND SCIENCE LOUISIANA TECH UNIVERSITY

March 2018
LOUISIANA TECH UNIVERSITY
THE GRADUATE SCHOOL

We hereby recommend that the dissertation prepared under our supervision by Grace Naomie Joseph, B.S. entitled Cost-effective use of Follow-up colonoscopies and Fecal Immunochemical Colorectal Cancer Screening be accepted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Computational Analysis and Modeling.

Supervisor of Dissertation Research

Head of Department

Department

Recommendation concurred in:

Advisory Committee

Approved:

Director of Graduate Studies

Dean of the Graduate School

Dean of the College
ABSTRACT

Colorectal cancer (CRC), if not detected early, can be costly and detrimental to one's health. Screening techniques for colorectal cancer can help find and remove adenomas in an effort to prevent future CRC. A colonoscopy has the potential of identifying CRC early as well as preventing the disease. Fecal Immunochemical Tests (FITs) is an initial non-invasive screening technique that identifies the presence of the disease, and needs to be followed by a more invasive test (like colonoscopy) if anything is detected. However, the optimal frequency for follow-up colonoscopy, and screening colonoscopy with FITs is unknown, and may vary based on findings from colonoscopy screening and patient age.

A partially observed Markov process (POMP) was used to simulate the effects of a follow-up colonoscopy, FITs, and a combination of colonoscopy and FITs on the development of CRC. The POMP uses adenoma and CRC growth models to calculate the probability of an individual having colorectal adenomas and CRC. Then, based on mortality, quality of life, and the costs associated with diagnosis, treatment, and surveillance of colorectal cancer, the overall costs and increase in quality adjusted life years are calculated for each follow-up colonoscopy, FIT and colonoscopy combination, and FIT scenario.

This study found that spending $100,000 for a quality adjusted life year (QALY) gained will produce one follow-up colonoscopy after screening at age 50 years. The
optimal follow-up is 8.5 years later which gives 84.0 QALYs gained/10,000 persons). No follow-up colonoscopy was cost-effective at the $50,000 and $75,000/QALY gained thresholds. The intervals were relatively insensitive to the findings at screening colonoscopy.

Combining screening techniques for CRC (colonoscopy and FIT) is cost-effective at all three thresholds analyzed. All screening ages tested were cost-effective with approximately an 8 year increase in QALYs from screening age 50 to screening age 75. As a result of the protocol of combining screening tests being extremely cost-effective with added benefit (QALYs gained), coupling the tests is highly recommend.
APPROVAL FOR SCHOLARLY DISSEMINATION

The author grants to the Prescott Memorial Library of Louisiana Tech University the right to reproduce, by appropriate methods, upon request, any or all portions of this Dissertation. It is understood that “proper request” consists of the agreement, on the part of the requesting party, that said reproduction is for his personal use and that subsequent reproduction will not occur without written approval of the author of this Dissertation. Further, any portions of the Dissertation used in books, papers, and other works must be appropriately referenced to this Dissertation.

Finally, the author of this Dissertation reserves the right to publish freely, in the literature, at any time, any or all portions of this Dissertation.

Author ________________________________

Date ________________________________
DEDICATION

To Kaejha, Mama, Julie and Tiffany.
# TABLE OF CONTENTS

ABSTRACT iii

DEDICATION vi

LIST OF FIGURES xi

LIST OF TABLES xiii

ACKNOWLEDGMENTS xiv

CHAPTER 1 INTRODUCTION 1

1.1 Colorectal Cancer Background 1

1.2 Overview of Cancer 2

1.3 Colon and Rectum 3

1.4 Colorectal Cancer 4

1.5 Risk Factors of Colorectal Cancer 5

1.5.1 Uncontrollable CRC Risk Factors 6

1.5.1.1 Age 6

1.5.1.2 Family History of Cancer 6

1.5.1.3 Personal History of Inflammatory Bowel Disease 7

1.5.2 Controllable CRC Factors 7

1.6 Symptoms of CRC 7

1.7 CRC Stages 8

1.8 CRC Treatments 9

1.8.1 Local Treatments 10

1.8.2 Systemic Treatments 11
3.3.6.2 Discounting 39
3.3.6.3 Quality-Adjusted Life-Years (QALYs) 40

3.4 Cost-effectiveness Analysis 42
3.5 Sensitivity Analysis 43

CHAPTER 4 ANALYSIS OF SIMULATION PROTOCOLS FOR COLONOSCOPY FOLLOW-UP 45

4.1 Introduction 45

4.2 Follow-up Colonoscopy based on Patient Age and Screening 45
  4.2.1 Age 50 45
  4.2.2 Age 55 49

4.3 Cost-effective Interval that Maximizes Patient’s QALYs Gained by Age and Screening Results 52
  4.3.1 Results for Control Group 1 52
  4.3.2 Age 50 53
  4.3.3 Age of final colonoscopy 56

4.4 Sensitivity Analysis 57
  4.4.1 Low Screening Colonoscopy Costs 59
  4.4.2 High Cancer Cost 62
  4.4.3 Zero Percent Discount Rate 63

CHAPTER 5 COST-EFFECTIVENESS OF FECAL IMMUNOCHEMICAL TESTS (FITs) 64

5.1 Introduction 64

5.2 Overview of simulation process for FIT 65
  5.2.1 POMP Model for FIT 65
  5.2.2 FIT Protocol 66
5.2.3 Costs
5.2.4 QALYs
5.2.5 Cost-effectiveness Analysis
  5.2.5.1 Control Group: No screening from age 0 until death or age 100
5.2.6 Results from FIT only Protocol
5.2.7 Combination of Colonoscopy and FIT Tests

CHAPTER 6 CONCLUSIONS AND FUTURE WORK
6.1 Conclusions
6.2 Future Work
6.3 Limitations of the study

APPENDIX A MATLAB CODE USED FOR MODELING
A.1 Overview

Bibliography
LIST OF FIGURES

Figure 11: Diagram of the colon and rectum. 4

Figure 12: Diagram showing healthy colon vs unhealthy colon. 5

Figure 13: Diagram showing before and after a resection surgery of the colon 10

Figure 21: Instrument used to perform a colonoscopy. 14

Figure 22: Instrument used to perform a sigmoidoscopy. 15

Figure 23: Difference in colonoscopy procedure (yellow) and sigmoidoscopy (green). 16

Figure 31: General overview of simulation process. 27

Figure 32: Adenoma growth model: Adenoma growth, CRC growth, Symptoms. 28

Figure 33: Model fit of overall CRC mortality using breakpoint model. 32

Figure 34: All-cause mortality data. 34

Figure 41: Efficiency vs Surveillance intervals at age 50 for no neoplasia finding. 46

Figure 42: Efficiency vs Surveillance intervals at age 50 for 1 to 2 polyps 47

Figure 43: Efficiency vs Surveillance intervals at age 50 for 3 plus polyps 47

Figure 44: Efficiency vs Surveillance intervals at age 50 for 1 to 2 advanced polyps only finding. 48

Figure 45: Efficiency vs Surveillance intervals at age 50 for 3 plus advanced polyps only finding. 48

Figure 46: Efficiency vs Surveillance intervals at age 55 for a general all screening results- an average of all the findings combined. 50

Figure 47: Efficiency vs Surveillance intervals at age 60 for a general all screening results- an average of all the findings combined. 50
**Figure 48:** Efficiency vs Surveillance intervals at age 65 for a general all screening results- an average of all the findings combined.

**Figure 49:** Efficiency vs Surveillance intervals at age 70 for a general all screening results- an average of all the findings combined.

**Figure 410:** Efficiency vs Surveillance intervals at age 75 for a general all screening results- an average of all the findings combined.

**Figure 411:** Effectiveness vs Surveillance intervals (ΔQALY) at age 50 with no neoplasia.

**Figure 412:** Effectiveness vs Surveillance (ΔQALY) at age 50 with 1 to 2 small polyps only.

**Figure 413:** Effectiveness vs Surveillance (ΔQALY) at age 50 with 3 plus small polyps only.

**Figure 414:** Effectiveness vs Surveillance (ΔQALY) at age 50 with 5 plus small polyps only.

**Figure 415:** Effectiveness vs Surveillance (ΔQALY) at age 50 with 3 plus advanced polyps only.

**Figure 416:** Efficiency vs Surveillance intervals at age 50 with all screening results.

**Figure 417:** Efficiency vs Surveillance at age 55 with all screening results.

**Figure 418:** Efficiency vs Surveillance at age 60 with all screening results.

**Figure 419:** Efficiency vs Surveillance at age 65 with all screening results.

**Figure 51:** Overview of FIT methods.

**Figure 52:** QALYs gained for the best protocol for FIT only.

**Figure 53:** Efficiency Vs Age for the best FIT only protocol.
LIST OF TABLES

Table 31: Annual transition probabilities between adenoma and CRC states. 29
Table 32: Annual transition probabilities between adenoma and CRC states. 31
Table 33: Annual transition probabilities between adenoma and CRC states. 33
Table 34: Costs used in the cost-effectiveness analysis. 37
Table 35: QALYs used in the cost-effective analysis. 41
Table 41: Ranges of follow-up colonoscopies that were cost effective at age 50 for all screening findings. 49
Table 42: Inter-colonoscopy intervals that maximize QALYS gained for the $100,000 threshold. 56
Table 44: Inter-colonoscopy intervals for sensitivity analyses. 57
Table 51: FIT only Analysis. 69
Table 52: Combination of FIT and colonoscopy analysis with control group 1. 72
Table 53: Combination of FIT and colonoscopy analysis with control group 2. 72
ACKNOWLEDGMENTS

I would like to start off by thanking God for giving me the strength and patience to take on this degree. I would also like to thank him for giving me such a great family structure to fall on for support and guidance in every way possible.

To the greatest advisor any student can ask for, Dr. Eric Sherer, I would like to express my sincere gratitude and appreciation to you for taking me under your wings. I will admit I started off a bit raw but with your profound knowledge and sincere nature, I turned out pretty okay. Thank you for all the support that you provided throughout my time here and for understanding me.

I would like to thank Dr. Jinko Kanno, Dr. Sumeet Dua, Dr. Prema Dua and Dr. Pedro Derosa for serving on my dissertation committee. Also, I would like to say a special thank you to Dr. Pedro Derosa and Dr. Bala Ramachandran, for introducing me to Dr. Sherer when I was in need of an advisor.

To my daughter, Kaejha Stewart, thank you for being the most loving, kind and sincere kid anyone could ask for. I will admit pursuing this degree and raising you to be a beautiful, intelligent, strong and powerful little lady was a bit daunting at times but I wouldn’t trade it for anything. You taught me so much there is to know about life, both ups and downs, and it was all well worth it.

I would like to thank my mom, Maryann Joseph, for being one of the main and biggest support systems throughout my journey. Though your education was cut short at
an early age due to you having to raise your siblings as a result of your parents’ death, you ensured that your children received the help and guidance they needed to move way further than you did. For that I will forever be thankful. Thank you for all the prayers every minute, every second, and every day.

   To my sister, Julie Joseph, thank you for believing in me and pushing me beyond my limits, and for all the support in all ways. The constant prayers and kind words to always push through even when times were tough really helped. You are truly an inspiration and the greatest big sister anyone can ask for.

   I would like to thank my best friend, Tiffany Victor, for the constant support throughout our shared journey. We started this journey together and supported and encouraged each other throughout. On the days where I felt I had no idea what I was doing or why I was doing this, you constantly reminded me and encouraged me.

   To my close friends Kevin Stewart, Kahla St. Marthe, Jermaine Morris, Germaine Pierre, Kahleen St. Marthe, Kerneisha Roachford, and Kazim Abbott, who helped in whatever way they could with anything I needed, I thank you from the bottom of my heart. I would like to say a big thank you to Jonas Montoute for helping me out with the formatting of this dissertation which was not much fun. He sacrificed himself right before me and passed on the good details and advice to make it easier for me. Also, I would like to thank Farid Heidarinejad for all his support and help.

   Lastly, I would to thank everyone who had some impact on this journey, whether it was big or small. To Christine Davis, Nicketta Stewart, the Tamakloes, and Gretchen Jones, a big thank you!
CHAPTER 1
INTRODUCTION

1.1 Colorectal Cancer Background

Colorectal cancer (CRC), if not detected early, can be costly and detrimental to one's health. Colorectal cancer is the third most commonly diagnosed cancer, and also the third leading cause of cancer-related deaths in both men and women in the United States [1]. It is estimated that in 2017, the US will record 135,430 new cases (95,520 new cases of colon cancer and 39,910 new cases of rectal cancer) [2]. For men, the lifetime risk for developing the disease is about 4.7%, and for women it is about 4.4% [1]. The disease is expected to cause 50,260 deaths in 2017.

The death and incidence rates for colorectal cancer over decades have been declining. This could be due to several reasons including, but not limited to, the fact that screening is allowing more colorectal adenomas to be found and removed before they develop into cancers, and also more cancers are being found earlier when the disease is easier to cure.

Colorectal cancer, however, is both treatable and preventable when detected at an early stage. About 90% of CRC cases develop in persons 50 years and older [3]. Some major risks factors of the disease include age and personal history of one or more colorectal adenomas. The majority of CRCs are believed to grow from adenomatous
polyps and removing the polyps has been shown to prevent CRC [4]. If CRC is detected early, it is more likely to be cured.

Appropriate screening exams can help find and remove adenomas in an effort to prevent future CRC. Several screening techniques have emerged over decades in order to help lower CRC. The two main approaches for screening CRC are (1) assessing the stool for blood or DNA and (2) visual examinations. Some common stool-based tests which are used are guaiac-based fecal occult blood test (gFOBT) and fecal immunochemical test (FIT) and some common visual-based tests are colonoscopy and flexible sigmoidoscopy.

The motivation of this dissertation is to model screening and surveillance strategies for colorectal cancer and to assess their cost-effectiveness.

1.2 Overview of Cancer

When cells in the body abnormally grow out of control and take over the normal cells, cancer results. It can start in any part of the body with no warning and can affect any individual [5]. Healthy cells grow, divide and die appropriately whereas cancer cells are likely to proliferate uncontrolled and outlive healthy cells. As these abnormal cells grow, they can spread away from their original location through the blood and lymph systems [6].

Some types of cancers can grow and spread quicker than others depending on the part of the body where they are located. As a result, different types of treatments are needed. These treatment options may include surgery and/or the use of chemotherapeutic drugs to either remove or reduce the diseased cells.

Some cancers sometimes form a lump which is called a tumor or a growth. However, not all lumps are cancerous [6]. Doctors are responsible for removing the
lumps and testing them to find out if they are cancerous or not. Lumps that are found to be non-cancerous are considered benign. They often times resemble normal body cells although sometimes they can grow very large enabling them to inhibit normal body functions but they don’t metastasize. Once benign tumors are removed they rarely return [6].

Lumps that are found to be cancerous are considered malignant and need to be dealt with as soon as they are detected. The malignant cells travel to other sites in the body and invade nearby tissue which can result in the growth and formation of new tumors. Cancers are named for the site where they began [6].

When cancers are detected it is better to start treatment sooner than later in order to increase life years.

1.3 Colon and Rectum

The colon and rectum are part of the digestive system; they make up the large intestine. The colon takes up most of the digestive system (5 feet) and the rectum occupies a small portion (6 inches). The colon has four main parts which are named by the direction the food is travelling [7]:

1. The ascending colon- extends upwards on the right side of the abdomen,
2. The traverse colon-travels across the body from right to left,
3. The descending colon- travels down on the left side of the abdomen,
4. The sigmoid colon- “S” shaped section that joins the rectum.

The ascending and traverse parts are referred to as the proximal colon collectively, and the descending and sigmoid parts are the distal colon. After the food goes through the small intestines, whatever is left, the colon absorbs salt and water from
it. Once the food goes through the colon, whatever is not needed goes to the rectum and is stored until it gets out of the body via the anus. **Figure 11** shows the digestive system with colon and rectum [7].

![Diagram of the colon and rectum](image)

**Figure 11:** Diagram of the colon and rectum[8].

### 1.4 Colorectal Cancer

Colorectal cancer is cancer that starts in the inner lining of the colon or the rectum. They can be named colon or rectal cancer independently based on where they start. According to the American Cancer Society, because these cancers have many similar characteristics they are often grouped together as colorectal cancer. The disease evolves from a growth called a polyp. These polyps can take approximately 10 to 20 years to develop into cancer but not all result in cancer [3].

The type of polyp determines the possibility of cancer transition. According to the American Cancer Society, there are two main types of polyps: adenomatous polyps (adenomas), and hyperplastic and inflammatory polyps. The adenomas can be
precancerous, but the hyperplasic and inflammatory polyps, even though they are more common, are not [9]. The size (> 1cm) and the number found (>2) of the polyp may increase the chances of being cancerous. Once the cancer cells are in the walls of the colon, they can grow into the blood or lymph vessels then to lymph nodes and eventually to other parts of the body [6].

Early detection of the disease is the most accurate method to reduce CRC-related death. Figure 12 shows the difference between a normal colon and an abnormal colon with cancer.

![Healthy Colon vs Colon Cancer](image)

**Figure 12:** Diagram showing healthy colon vs unhealthy colon[10].

### 1.5 Risk Factors of Colorectal Cancer

According to World Health Organization, any attribute, characteristic or exposure of an individual that increases the likelihood of developing a disease or injury is a risk factor. As with everything else, different cancers have different risks factors. Some risk
factors such as smoking can be controlled whereas familial history cannot. There is no indisputable evidence to prove what causes colorectal cancer but some researchers have found several risk factors that might increase a person’s chances of getting the disease. Some of these risk factors include family history, inherited syndromes, racial and ethnic background, diet, inactive lifestyle, smoking, alcohol use, age, history of colorectal cancer or polyps, history of inflammatory bowel disease (IBD), obesity, type II diabetes, etc. [11]. Both lifestyle and genetic influences play a role in the disease development.

1.5.1 Uncontrollable CRC Risk Factors

1.5.1.1 Age

The strongest risk factor is age. As people age, they have greater chances of obtaining the disease. Approximately 90% of CRC cases develop in persons 50 years and older [3]. That does not mean that CRC cannot occur at younger ages. It can actually occur in teens in very rare cases [12-15].

1.5.1.2 Family History of Cancer

A very strong predictor of developing colorectal cancer is associated with having a first degree relative with the disease. According to the American Cancer Society 1 in 5 people who develop CRC have other family members you have had it [16]. Having one or more first-degree relatives younger than 45 years with the disease, puts you at an increased risk for the disease [17],[18].

Having family members who have had adenomatous polyps (they can become cancerous) have also been associated with increased risk of CRC. Someone with a family history of adenomatous polyps or colorectal cancer is more prone to developing the disease than others with no history [16].
1.5.1.3 **Personal History of Inflammatory Bowel Disease**

Any inflamed bowel disease (IBD) such as Ulcerative colitis and Crohn’s disease increases the risk of CRC [19],[20]. IBD occurs when the colon is inflamed over an extended period of time. If an individual has IBD for many years, dysplasia often develops. Dysplasia describes the cells in the lining of the colon or rectum that may seem abnormal but are not cancerous. They however can evolve into cancer with time [1], [16].

1.5.2 **Controllable CRC Factors**

Some other possible factors that may increase the risk of developing CRC, which can be controlled by an individual, include habits such as smoking and drinking. Also having diets with foods high in fat and cholesterol and low in fiber can deteriorate your general health quicker [16]. Some studies even link increase CRC risk with working night shifts regularly [21].

1.6 **Symptoms of CRC**

It may take a while for CRC symptoms to appear. The symptoms depend on the stage of the cancer and where it was formed or where it is currently located. Some common symptoms associated with the disease include[1]:

- Change in bowel habits – diarrhea, constipation, narrowing of stool,
- Rectal bleeding,
- Cramping or abdominal pain,
- Weight loss and
- Weakness and fatigue.

These symptoms can be caused as a result of a different condition. Most of these symptoms are most likely due to the disease being spread to other parts of the body; the
later stage of the disease. This is why it is recommended to get screening early to possibly reduce or avoid the full development of the disease.

The symptoms of CRC depend on the location of the cancer. Cancer on the right side of the colon may go unnoticed for a very long time. People may feel a mass; have abdominal pain, fever, profuse sweating (especially at night) and anemia. If the cancer is on the left side, the main symptom that is observed is a change in bowel habits. They also present similar symptoms like abdominal pain, abdominal distention, vomiting and constipation similar to right side cancers.

1.7 CRC Stages

The extent of the cancer in the body is grouped by stages [1]. A stage is a crucial factor in deciding how to treat the disease. After CRC has been diagnosed, experts use different characteristics to determine the cancer’s stage. They typically look into how far the cancer has grown in the walls of the intestine, whether it has reached any nearby structures and also if it has spread to other distant organs [1].

The data collected after diagnosis is very important to cancer registries and hospitals. Based on what their needs are, they use various staging systems [22]. Some of the major systems used to group cancer stages are (1) The American Joint Committee on Cancer’s TNM System [23] (2) Surveillance, Epidemiology, and End Results (SEER) Extent of Disease (EOD) and (3) SEER Summary Staging Guide [24].

The SEER Summary Staging Guide was published in 1977 [22] and it was more site specific than the broad staging categories that were available at the time. The stage categories were defined as (1) In situ – a neoplasm is present and has all the characteristics of malignancy with the only exception that it is not invasive, (2) Localized
– an invasive malignant neoplasm is present and it is confined entirely to the organ of origin, (3) Regional – a malignant neoplasm is present and has gone past the limits of the organ of origin into nearby organs or tissues or via the lymphatic system to regional lymph nodes and (4) Distant - a malignant neoplasm is present and it has spread to other parts of the body isolated from the primary tumor to distant organs, tissues or via the lymphatic system to distant lymph nodes [22]. The five-year survival of the disease is 90% if it is diagnosed in the localized stage, 68% for the regional stage and only 10% for the distant stage [25].

The staging system most often used is The American Joint Committee on Cancer’s TNM System [1]. This system is based on three important pieces of information:

1. The extent of the primary tumor (T) into the wall of the intestines and other nearby areas,
2. Whether there is a spread to nearby lymph nodes (N),
3. Whether the cancer has metastasized (M) to other organs of the body.

These letters (TNM) are normally followed by numbers or letters to give more information about each of the factors. The greater the number the more advance the cancer.

1.8  **CRC Treatments**

Several treatment options exist for colorectal cancer based on type and stage of the disease. American Cancer Society groups CRC treatments into two categories: local and systemic [16].

1.8.1  **Local Treatments**
Local treatments tend to treat just the tumor without affecting any other part of the body and are more likely it be used for earlier stage CRC. Some local treatments include surgery, radiation therapy, and ablation or embolization.

- Surgery is the primary treatment when the cancer is in its early stage. The surgery is chosen based on the stage of the cancer and its location. Some surgeries include polypectomy and local excision, colectomy, diverting colostomy, local trans anal resection, trans anal endoscopic microsurgery (TEM), lower anterior resection (LAR), and abdominoperineal resection (APR) [1],[26]. A resection surgery removes integrated sections of the colon as seen in Figure 13.

![Figure 13: Diagram showing before and after a resection surgery of the colon [27].](image)

- Radiation therapy focuses on high-energy rays to destroy cancer cells. It is sometimes paired with another treatment option (Chemotherapy) to be more effective [16],[28].
• Ablation deals with treatments that destroy the tumors without having to get rid of them. Embolization deals with injecting substances to try to reduce blood flow to cancer cells in the liver. The two treatment options can be combined [16],[29].

1.8.2 Systemic Treatments

Drugs can also aid in the treatment of CRC. They can be given orally or directly into the bloodstream. Some of these treatments include chemotherapy, targeted therapy, and immunotherapy [30],[16].
CHAPTER 2
CRC SCREENING AND SURVEILLANCE

2.1 Background

Recently computer models have been used to better identify those at risk for colorectal cancer and to help determine the best treatment for them. With the help of effective computer models, the progress and effect of CRC in terms of cost and life-years gained for screening and surveillance methods can be predicted. Without these models it would be impractical and very expensive to study the effects of the disease on people.

Colorectal cancer is a disease suitable for screening [31],[32]. Screening tests can assist in finding pre-cancerous adenomatous polyps, which can then be removed before becoming cancerous. CRC screening tests also aide in finding colorectal cancer early when it is best treated.

2.2 Screening

Colorectal cancer is both treatable and preventable when detected at an early stage. Appropriate screening exams can help find and remove adenomas in an effort to prevent future CRC. Several screening techniques have emerged over the past few decades to help lower CRC [33-36]. Screening methods for CRC differ in performance features, complication rates, suitability and costs [36-39].
Screening tests can be compared based on accuracy by how specific or sensitive they are. Sensitivity and specificity are mostly used to determine the effectiveness and quality of the test. Sensitivity measures the proportion of positives that are found in a sick population and is the capability to identify true positives. If a test records few false negatives and several true positive it is considered highly sensitive. Mathematically, sensitivity ($S$) is expressed as [40]

\[ S = \frac{TP}{TP + FN} \]  

or

\[ S = \frac{TP}{TP + FN} \]

Specificity ($Sp$) measures the number of negatives that are accurately identified as such. If few false positives and many true negatives are generated for a screening module then that module is considered to have high specificity. Mathematically, specificity is expressed as

\[ Sp = \frac{TN}{TN + FP} \]

or

\[ Sp = \frac{TN}{TN + FP} \]

There are two main recognized approaches for CRC screening namely endoscopic tests (visual examinations) and non-endoscopic tests (the stool is assessed for blood or DNA).

2.2.1 Endoscopic Test
Endoscopic tests allows doctors to look at the inside lining of the organ. The two main endoscopy tests involved in CRC screening is sigmoidoscopy and colonoscopy.

- Colonoscopy- screening with the removal of adenomas is an effective strategy for reducing colorectal cancer incidence and mortality [41]. During a colonoscopy, a specialist uses a colonoscope (see Figure 21), which is a long, thin and flexible tube with a camera and light on the end, and inserts it in the rectum going through the entire colon. In Figure 21, A is the position of the colonoscope; B is the endoscopic view; and C is the detail of the colonoscope tip.

![Figure 21](image)

**Figure 21**: Instrument used to perform a colonoscopy [42].

- Sigmoidoscopy- Screening with flexible sigmoidoscopy has shown to reduce the rates of CRC diagnoses and deaths compared to no screening [43-45]. During
screening, the doctor examines only a portion of the colon and rectum using the same tool described for the colonoscopy. (See Figure 22). Sigmoidoscopy is often easier on patients but can miss polyps if they are in the distal colon [43].

![Figure 22: Instrument used to perform a sigmoidoscopy [46].](image)

During these two procedures if some abnormal colon issue is found, it is collected and sampled. This is called a biopsy. If polyps are found during the colonoscopy or sigmoidoscopy then a polypectomy is performed. The polyp is captured and cut from the colon lining using a wire. Patients prepare for the procedures by removing all stools from the colon in order to improve visibility. The procedures are almost always performed as outpatient surgeries. Figure 23 shows a comparison between a colonoscopy and a sigmoidoscopy.
2.2.2 Non-endoscopic Test

Fecal Immunochemical Tests (FIT) and Guaiac Fecal Occult Blood Tests (gFOBT)

Stool based screening tests are low-cost, non-invasive and is easily accessible for patients [48]. Fecal tests look for hidden blood in the stool. This could indicate an early sign of CRC. FIT and gFOBT are initial screening tests for CRC. It has been shown that testing for blood in the stool significantly reduces the incidence and death rate for CRC [48-51]. Some proof that the fecal blood tests reduce mortality was provided in some randomized trials around the 1980’s [52]. CRC mortality was reduced by 15% to 33% when positive results were followed by a colonoscopy [49-51, 53].

Figure 23: Difference in colonoscopy procedure (yellow) and sigmoidoscopy (green)[47].
A gFOBT uses the chemical guaiac to detect heme in the stool. Heme is an iron-containing compound of the porphyrin class that forms the non-protein part of hemoglobin and some other biological molecules [54]. During the process, hydrogen peroxide is dropped on the paper containing the stool and if blood is present, the paper will change to a blue color. The test is done with a take home kit so patients can have their privacy. From randomized trials [49-51, 53], the sensitivity of gFOBT ranges widely from 30%-90% [36], and its specificity around 78% [55]. FIT uses antibodies to detect human hemoglobin protein in the stool. The FIT procedure includes as supply kit with a collection bottle, absorbent paper, biodegradable collection paper, and a sample-mailing envelope. The kit will come with instructions on how to collect the sample. For the FIT test, only one sample needs to be collected and there are no dietary restrictions unlike the gFOBT.

The two tests are similar but FITs uses a more accurate technology in detecting blood in the stool. For that reason, studies [36], [55] show that the sensitivity of FIT is between 75% -90% with specificity about 90%. FIT has a higher sensitivity for screening relevant tumors, early stage cancers and advanced adenomas than gFOBT [36, 56, 57].

Since fecal tests have poor sensitivity to premalignant lesions, they must be repeated more often usually every 1-2 years [48]. FIT is now considered a better alternative than gFOBT because of the better sensitivity. Since the interval for receiving fecal tests is shorter due to sensitivity and the costs of fecal tests are relatively low, combining fecal tests and with other screening techniques like sigmoidoscopy and colonoscopy may actually be more effective. Studies [58],[59],[60] have evaluated combined screening techniques. These studies focused on fecal blood test in conjunction
with sigmoidoscopy and found out that the likelihood of early detection as well as survival rates of CRC increased.

### 2.3 Complications due to Screening

There are some adverse effects associated with the procedures of colonoscopy. Some include, but are not limited to, death, bowel perforation, bleeding, serosal burn, reaction to sedation medicine, post-polypectomy hemorrhage, bleeding with transfusion and so on. The chances of having serious adverse effects of colonoscopy after screening are low but with increasing age they are relatively higher for follow-up colonoscopies [61].

### 2.4 Surveillance

Based on what was found at screening, a surveillance regimen is followed. People with prior history of colorectal adenomas or CRC should be checked more often than if nothing was found. This is because the risk of new colorectal adenomas or CRC forming is greater so a more aggressive approach needs to be implemented for such individuals. Surveillance intervals are typically shorter and it depends on the size and number of adenomas found.

### 2.5 Current Guidelines for Screening and Surveillance of Colonoscopy

Although in the last decade there has been an increase in the number of technologies available for CRC screening, colonoscopy screening rates are increasing while other technology rates are either declining or remain constant [62]. According to a study [63] done to determine whether Medicare reimbursement policy changes were associated with an increase in either colonoscopy use or early stage CRC diagnosis,
colonoscopy use increased by 5% each quarter during the study. The study showed that between 1993 and 2002, there was a staggering six fold increase in colonoscopy usage for detecting CRC.

Colonoscopy is thorough because it goes through the entire colon and also it is very sensitive. About 95% of cancers are detected during a colonoscopy. The National Polyp Study (NPS) found that patients who underwent a colonoscopy removing adenomas had a 76% to 90% lower incidence of CRC relative to patients without polypectomy [64]. The results of the National Polyp Study [64] revealed that colonoscopy polypectomy resulted in lower-than-expected incidence of colorectal cancer which supports the claim that adenomas grow to adenocarcinomas and also support the effectiveness of the current practice of identifying and removing adenomatous polyps [3]. As a result, recent trends in CRC incidence and mortality reveal declining rates, which have been attributed to reduce exposure to risk factors, screening effect on early detection and prevention through colonoscopy polypectomy, and improved treatments [3]. As colonoscopy rates increase, several specialty societies support colonoscopy as the ideal screening method for CRC [3].

The Microsimulation Screening Analysis (MISCAN-Colon) and Simulation Model of CRC (SimCRC) models are two models that compare strategies for screening that vary by age [65],[66],[67].

- MISCAN-Colon

This simulation model [67] was developed to examine trends in CRC due to various changes in lifestyle, improvement of treatment and application of screening approaches. It uses a Monte Carlo method to simulate all the events such as birth, death,
adenoma incidence, and transitions from one state of the disease to the other. It consists of three parts: (1) demography, (2) natural history, and (3) screening. The model combines demographic and epidemiological information from the SEER program. Individual life histories are generated from demography to form a population. Each patient consists of a date of birth and a date of death from other causes other than CRC. The model simulates a population of people where CRC and adenomatous polyps developed leading to a clinical diagnosis, treatment, and possibly death. In the natural history part, fictitious individual life histories where numerous colorectal lesions can merge are generated. Using the adenoma-carcinoma sequence of Vogelstein [68] and Morson [69], the natural history model was developed. In the screening part, some life histories will be changed. Age of screening, length of screening, the screening test and the diagnostic follow-up scheduled after a positive test is included in a screening policy. After each part the life histories are changed. The screening and natural history part are run simultaneously.

- SimCRC

The SimCRC [65] is a population-based microsimulation model that was developed in a way that it can simulate the US population from 1970 to a future year and it can also run a single birth cohort. The model is based on a state-transition Markov model designed to evaluate cost-effectiveness of screening [33]. The model consists of a natural history part, a screening part and a treatment part. The natural history follows the adenoma to carcinoma sequence as a function of age, sex, race, and risk factors. With the screening part, adenomas are detected and removed and early diagnosis of preclinical
CRC is possible. Once you have been diagnosed with CRC, then the treatment part is implemented.

### 2.6 Screening Age

When the disease is detected early it can be prevented thus reducing mortality rates but there is no ideal age to begin colonoscopy or to determine how often it should be done. The risk of the disease increases with age so beginning screening later will have more CRC develop while starting screening too early will likely not prevent as many cancers due to fewer adenomas having been developed. As each individual age, there is a chance that an adenomatous polyp forms and grows. One or more adenomas can happen in someone and each can develop independently into CRC [65]. The microsimulation models (MISCAN-Colon and SIMCRC) simulate colorectal disease progression in a large population of individuals from birth to death [65], [66]. These models are very useful because they allow for continued investigation of the results of varying screening age. Simulation models are sometimes used to determine if there is a protocol that is superior to existing care. They can also be used to predict the life-years gained relative to no screening, the number of colonoscopies required, the age at which to begin screening, age at which to stop screening, and screening interval [65].

Based on such simulation models and clinical evidence, the US Preventive Task Force guidelines recommend that people at average risk of colorectal cancer be screened starting at age 50 and should end at age 75 [70], [66]. People with no symptoms of the disease or personal or family history are considered average risk. They also suggest that people at average risk should undergo a colonoscopy every 10 years [66]. However, there is no formal guidance for the frequency of follow up colonoscopy after a colonoscopy in
which adenomas were found. Such patients are at higher risk for subsequent CRC [71] so the inter-colonoscopy intervals are typically shorter than 10 years. In addition, post-colonoscopy risk increases with age, but the frequency with which follow-up colonoscopy based on colonoscopy history and age should be done is unknown. This frequency will likely vary based on the risk factors for CRC at follow up.

2.7 Cost-effectiveness Analysis

Since the incidence of CRC is age-dependent, the effectiveness and cost of colonoscopy screening depend on the age at which the first examination is performed [72]. Although it has been found that colonoscopy screening is effective [71], the effectiveness comes at a significant additional cost [33]. Possible explanations for the increased cost is the need for repeated colonoscopy screening and surveillance in patients found to have adenomas on initial screening [33]. Cost-effectiveness of screening methods is often used in CRC research to compare different testing methods [73]. Cost-effectiveness analysis is considered a reasonable tool that properly compares the health and economic costs of different techniques, in turn helping decision makers to identify those that will produce the most health profits, given their resources limitations [73].

With the costs of providing these different interventions combined with the benefits associated with them, a cost-utility ratio result. This is the proportion between the cost of a health-related intervention and the benefit it produces in terms of number of years lived. This gives the incremental cost per QALY gained [74]. With the high costs associated with colonoscopy screening, well designed reviews are needed to assist decision makers in recognizing the effectiveness of the procedure in terms of QALYs gained [75]. QALY take into consideration both the quantity and the quality of life
generated by healthcare interventions. They provide a common way to assess the extent of the benefits gained in terms of health related quality of life and survival of the patient [75]. A QALY places a weight on time in different health states. A year of perfect health is valued at 1 and a year of less than perfect health is worth less than 1. Death is equal to 0; there are some health states that are considered worse than death which results in negative scores [75].

Since it has been proved that CRC screening by colonoscopy is more effective compared to other strategies [76], the economic impact has been evaluated in several cost effectiveness analyses. The majority of these analyses are based on the US population [73]. A systematic review conducted by Lansdorp-Vogelaar [73] on the cost-effectiveness of CRC screening showed that the cost-effectiveness ratios for all established screening strategies compared to no screening were less than $50,000 per life-year gained [73]. For colonoscopy screening the estimates were between $9,038 and $22,012 per life-year gained. Half of the studies found colonoscopy to be the most effective strategy. There is some variability on what is the appropriate benchmark. After over 20 decades in the US, the ratio of $50,000 per QALY gained by using a health care intervention still acts as a yardstick for value of care in the health policy system [77]. We would assume that based on inflation, economic growth and changes the yardstick would change but the benchmark still persists. Some argue that $100,000 is a better benchmark, some even argue, on the assumption of people’s values and attitudes toward risk, that the threshold should be two to three times the per capita annual income, which based on the US per capita income of ~$54,000 would put the threshold at about $110,000 to $160,000 per QALY today [77]. Other societies even say a threshold of $200,000 to $300,000 per
QALY can work based on 1) the increases in health care expenditure over time and the health gains associated with those increases, 2) various tools to determine how much people are willing to pay for health gains or 3) the trade-offs people make in the workplace between pay and safety risks [78],[79]. Between 1990 and 2012, many authors have referenced different thresholds for US-Based Cost-Utility Analyses [77] with $50,000 and $100,000 per QALY still remaining the most widely used thresholds.

Several other cost-effectiveness analyses of CRC screening have been published and they generally confirm the findings of the systematic review that CRC screening is cost-effective compared to no screening but none of them has yet been found out to be most effective or have the most attractive cost-effectiveness ratio per QALY gained [73]. This is why in this study a cost effective analysis that will help to find the proper balance between life years gained and costs than what is currently recommended is performed. This work also determines the most favorable inter-colonoscopy interval.

### 2.8 Research Need

In the growing field of colorectal cancer, a needs area is determining the cost-effective inter-colonoscopy interval. While the MISCAN-Colon and SimCRC discussed earlier performed cost analysis of existing colorectal cancer screening protocols and evaluated the optimal age to receive a one-time colonoscopy, no work has been done to optimize the inter-colonoscopy interval. Because clinical studies have shown that post-colonoscopy CRC risk varies with age and prior colonoscopy results, it is highly likely that the cost-effective inter-colonoscopy intervals will differ from those of currently recommended for screening. Although colonoscopies are said to be effective [41], [76], the high price at which they come, can be a burden to most patients. Cost-effectiveness
analysis will not tell you what the best intervention is, but it will let you know which intervention will provide the utmost health profits, given the decision maker's willingness to pay for a unit of health [73].

Finding cost-effective use of screening procedures based on patient age and colonoscopy history will help determine the appropriate interval between colonoscopies. This is important because it can save costs. If the time interval is too long, then the risk of CRC increases leading to the need for more expensive treatments. Whereas, if the interval is too short, patients will receive unnecessary treatments and this comes with high costs.

Many studies have shown the different ages for colonoscopy screening that are more effective. In this decision analysis model [73], screening at age 40 was less cost-effective than screening at age 50. Another study found that screening at age 60 was more cost-effective than screening at age 50 or 70 [74]. The US Preventive Task Force guidelines recommend that people at average risk of colorectal cancer be screened starting at age 50 and should end at age 75 [65,], [66]. Unfortunately, currently research has not been done using a partially observed Markov process (POMP) to determine the most appropriate interval of performing a colonoscopy based on patient age and prior colonoscopy history. We intend to begin addressing this need to determine whether patients’ history or age determines how soon they return for a follow-up colonoscopy.

Chapter 3 seeks to address the following questions:

- Is surveillance colonoscopy based on age and screening findings cost-effective?
- What is the optimal interval that maximizes patient QALYs?
CHAPTER 3

COMPUTATIONAL TECHNIQUES

3.1 Overview

An overview of the simulation process is shown in Figure 31. A partially observed Markov process (POMP) was used to simulate the effects that intervening with different follow-up colonoscopy scenarios had on the development of CRC. The POMP uses an adenoma and CRC growth model to calculate the probability of a patient having colorectal adenomas; asymptomatic local, regional, or distant CRC; and symptomatic local, regional, or distant CRC as the patient ages. Some of these adenomas and CRCs are then detected and removed during colonoscopy with the detection rate depending on the size of the neoplasia. The natural history model is coupled to a mortality rate model that calculates the probability of dying due to natural causes or CRC related mortality. Then, based on mortality, quality of life, and the costs associated with diagnosis, treatment, and surveillance of colorectal cancer, the overall costs and benefits in terms of QALYs are calculated for each follow-up colonoscopy scenario to determine its effects for the cost-effectiveness analysis. A diagram of the simulation process is available in Figure 31.
3.2 Markov Process

One of the most common and important of random processes is a Markov process. They were named after Andrei Markov. A Markov process is thought to be a random stochastic process (a sequence of events in which the outcome at any stage depends on some probability) where the number of possible outcomes or states is finite, the outcome at any stage depends solely on the outcome of the previous stage, and the probabilities remain constant over time [80]. It can be viewed by saying that the future and past of the process are independent of each other when the present is known.

A Markov process gives a more accurate solution to the problem you are addressing. A Markov process was chosen over a Monte Carlo Simulation because due to the nature of the number of simulations that needed to be run. Monte Carlo simulations also give an approximation compared to a more precise solution such as the Markov process.
3.3 Breakdown of each Component Methods

In this section, a detailed overview of each component of the proposed model will be discussed.

3.3.1 Colonic Neoplasia Growth Model

The colonic neoplasia natural history progression model combined transition rates from two studies each focusing on different phases of colonic neoplasia development: adenoma growth and CRC growth (see Figure 32).

![Diagram of Colonic Neoplasia Growth Model]

**Figure 32:** Adenoma growth model: Adenoma growth [81]; CRC growth [67], Symptoms [82].

To model adenomas growth, Sherer *et al.* [81] used serial colonoscopy results to identify the transition rates for the series of transition from diminutive adenoma (<5 mm) to medium adenoma (6-9 mm) to large/advanced adenoma (>10 mm) to CRC (Figure 32). It was assumed that multiple colorectal neoplasia can exist and grow/regress and that both the growth of each neoplasia and the appearance of new adenomas are independent of the other neoplasia. They tested whether each rate was age dependent and found that the rate of appearance of new adenomas varied with patients’ age but the transitions to more advanced neoplasia were age-independent. The rates obtained from Sherer *et al.* [81] were integrated over a short, one month interval to get the monthly transition
probabilities and the monthly transition probabilities were combined to get the yearly transition probabilities (see Table 31).

**Table 31:** Annual transition probabilities between adenoma and CRC states.

<table>
<thead>
<tr>
<th>Adenoma growth transition</th>
<th>Yearly transition probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diminutive to no adenoma</td>
<td>0.0181</td>
</tr>
<tr>
<td>Diminutive to medium adenoma</td>
<td>0.0163</td>
</tr>
<tr>
<td>Medium to diminutive adenoma</td>
<td>0.0716</td>
</tr>
<tr>
<td>Medium to large adenoma</td>
<td>0.0377</td>
</tr>
<tr>
<td>Medium adenoma to CRC</td>
<td>0.0025</td>
</tr>
<tr>
<td>Large to medium adenoma</td>
<td>0.0047</td>
</tr>
<tr>
<td>Large adenoma to local CRC</td>
<td>0.0028</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CRC growth transition</th>
<th>Yearly transition probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local CRC to regional CRC</td>
<td>0.22</td>
</tr>
<tr>
<td>Regional CRC to distant CRC</td>
<td>0.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CRC symptom transition</th>
<th>Yearly transition probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asymptomatic local CRC to symptomatic local CRC</td>
<td>0.17</td>
</tr>
<tr>
<td>Asymptomatic regional CRC to symptomatic regional CRC</td>
<td>0.22</td>
</tr>
<tr>
<td>Asymptomatic distant CRC to symptomatic distant CRC</td>
<td>0.50</td>
</tr>
</tbody>
</table>
To model CRC growth, the MISCAN-COLON model used SEER CRC prevalence data to identify the yearly transition probabilities between CRC stages and the probability of symptoms developing with CRC [67]. As people age, symptoms can develop as well. In the model, there are also transitions from asymptomatic to symptomatic states (which affects whether a patient receives a colonoscopy for symptoms and the neoplasia are detected) but we assume that the presence of symptoms does not affect the natural history of the neoplasia. A transition rate from asymptomatic to symptomatic [82] was used to determine probability of developing symptoms (see Table 31).

Each combination of adenomas and CRC defines a possible state of the colon. For example, the state, $s$, is given by number of diminutive adenomas, number of medium adenomas, and etc. and the probability that a patient is in the state at age $t$ is $P(s|t)$. After a time period, $t$, a patient in the state can transition to any other possible state (including remaining in the state) with a probability. Assuming no interventions, the dynamics of colonic state probability vector is the Markov Process and is given by

$$
\text{Eq. 31}
$$

where, initially, there are no neoplasia. When, a person ages, the risk of CRC increases. The model mimics this behavioral pattern: each adenoma progresses naturally towards CRC and eventually death.

3.3.2 Mortality Rate Model
Patient death rates due to both CRC-related mortality and all-cause mortality were considered in the POMP model. Mathematical models for the rates of both processes were developed.

3.3.2.1 CRC-related Mortality

The CRC-related mortality rate was simultaneously fit to two data sources: the five year survival data for local, regional and distant CRC for white males diagnosed in 2003-2009 all followed through 2010 (Table 32) and the annual overall CRC survival rate for 10 years after diagnosis for patients diagnosed in 2003 (Figure 33). A least-squares objective function was used to fit every data point from both data sources: 4 data points for 5-year survival (one for each stage including an unknown stage) and 11 data points for overall CRC survival rate (annually for ten years after diagnosis).

Table 32: Annual transition probabilities between adenoma and CRC states.

<table>
<thead>
<tr>
<th>CRC stage</th>
<th>Prevalence</th>
<th>CDC data for 5-year survival rate</th>
<th>Model-predicted 5-year survival rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local</td>
<td>44%</td>
<td>88.2%</td>
<td>88.1%</td>
</tr>
<tr>
<td>Regional</td>
<td>34%</td>
<td>70.1%</td>
<td>70.0%</td>
</tr>
<tr>
<td>Distant</td>
<td>18%</td>
<td>12.3%</td>
<td>12.2%</td>
</tr>
<tr>
<td>Unknown</td>
<td>5%</td>
<td>43.1%</td>
<td>43.1%</td>
</tr>
</tbody>
</table>
**Figure 33:** Model fit of overall CRC mortality using breakpoint model.

The CRC-related mortality rate varies by CRC stage so a different mortality rate was used for each CRC stage. The probability of surviving for a period of time after being diagnosed with the stage of CRC (where is either local, regional, distant, or unknown), is described by

\[ \text{Eq. 32} \]

where \( m \) is the mortality rate.

With those probabilities, the overall survival probability of CRC patients with time, \( t \), was calculated by combining the survival rates for each stage weighted by the prevalence of each stage and it is given by
where \( rCRC \) is regional CRC, \( lCRC \) is local CRC, \( dCRC \) is distant CRC, \( uCRC \) is the unknown CRC and subscript \( s \) represents survivor.

A constant mortality rate was not a good fit to the data because there was an over prediction in the earlier years since diagnosis and an under prediction in the latter years. A single breakpoint model was then applied on the mortality rates of each stage of the disease and provided an accurate fit to the data (Figure 419 and Table 33). The mortality rate increased as the CRC spread and the decrease in the mortality rate after a few years matches the expectation that the mortality rate is highest immediately after CRC diagnosis. In addition, the breakpoint was early as the cancer became more distant (7.3, 5.3, and 1.6 years for local, regional, and distant CRC, respectively).

**Table 33:** Annual transition probabilities between adenoma and CRC states.

<table>
<thead>
<tr>
<th>CRC stage</th>
<th>Initial mortality rate (1/year)</th>
<th>Breakpoint (years)</th>
<th>Second mortality rate (1/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local</td>
<td>0.025</td>
<td>7.3</td>
<td>0.018</td>
</tr>
<tr>
<td>Regional</td>
<td>0.068</td>
<td>5.3</td>
<td>0.025</td>
</tr>
<tr>
<td>Distant</td>
<td>0.600</td>
<td>1.6</td>
<td>0.085</td>
</tr>
<tr>
<td>Unknown</td>
<td>0.155</td>
<td>4.8</td>
<td>2.24286E-06</td>
</tr>
</tbody>
</table>

### 3.3.3 All-cause Mortality

The all-cause mortality data was obtained from Table III in the National Vital Statistics (NVS) - Life table database, 2010[83] (see Figure 34). The rates from the NVS are based per 100,000 in specified age groups. The midpoint of each age group was
obtained and plotted against the death rates from the NVS – life table database. A 6th order polynomial was used to fit with the existing National Vital Statistics data.

**Figure 34:** All-cause mortality data.

### 3.3.4 Partially Observed Markov Process (POMP)

Leading to the colonoscopy screening, the adenoma growth and mortality rate models are applied to the patients. A patient can die from natural causes, remain in a healthy state, develop symptoms and develop polyps or CRC. These patients are initially not in the patient pool because each hypothetical patient is considered healthy until their screening colonoscopy. To mimic the clinical classification of colonoscopy results, the probabilities of every possible combination of colonic neoplasia were used to calculated the probability of eight colonic neoplasia states: (1) no adenomas, (2) 1-2 non-advance adenomas only, (3) 3+ non-advanced adenomas only, (4) 1-2 adenomas with some
large/advanced adenomas, (5) 3+ adenomas with some large/advanced adenomas, (6) local CRC, (7) regional CRC, and (8) distant CRC.

A POMP was used to obtain the CRC predictions. Since colonoscopy is not 100% sensitive [81] the actual state is only partially observed. Sherer et al. [81] reported colonoscopies are more likely to be sensitive for larger adenomas (95.8%) as opposed to diminutive ones (39%). Several studies [39, 67, 81, 84-86] have shown that smaller adenomas are more commonly missed as opposed to larger ones. Sherer et al., in their study saw differences in patient’s age for diminutive, small, large and cancerous neoplasia.

Based on the findings at the screening colonoscopy, the patients will follow a follow-up colonoscopy regimen suitable to the findings obtained. Patients can die from natural death at any point in the model and are removed from the pool of patients. The probabilities collected from the POMP are incorporated into the costs and QALYS results to perform the cost-effectiveness analysis.

3.3.5 Follow-up Colonoscopy Scenarios

Different follow-up colonoscopy protocols were applied and the associated costs and benefits (in terms of QALYs gained) calculated. One follow-up colonoscopy was evaluated to determine its effects. Scenarios were evaluated for screening colonoscopy at ages 50, 55, 60, 65, 70 or 75 years, and a single follow-up colonoscopy from 2 years until 20 years in increments of 2 years. The CRC predictions were applied to the follow-up colonoscopy scenarios and the probabilities collected were incorporated into the costs and QALYS results. The cost per QALY gained was calculated (section 3.3.6.1) and the follow-up intervals were identified by quadratic interpolation.
3.3.5.1 Control Group

Control groups are essential parts to many research designs, they allow for meaningful comparisons to be drawn. They allow researchers to minimize the effect of all variables except the independent variable.

3.3.5.1.1 Control Group: Screening only

To measure the effect of follow-up colonoscopy a control group who received a screening but no follow-up colonoscopy was used. The costs and QALYs were obtained for patients who received surveillance colonoscopies vs. those who just had a screening and no surveillance. This scenario will give the optimal intervention at each screening age.

3.3.6 Outcomes of Partially Observed Markov Process

3.3.6.1 Costs

Costs are a main factor in a cost-effective analysis. The first set of costs that is typically observed with a disease is the cost of treating the disease like hospitalization costs, medicine costs, and doctor costs. Other costs include costs of screening procedures and new treatments, cost of missing work, transportation and travel costs, and so on.

Two types of costs were considered: costs associated with CRC treatment and costs associated with receiving colonoscopy [87-91] (see Table 34). We assumed all other costs to be equal between patients so there would be no cost difference based on colonoscopy.

Table 34: Costs used in the cost-effectiveness analysis.

<table>
<thead>
<tr>
<th>Type of treatment</th>
<th>Costs (Range for Sensitivity Analysis)</th>
</tr>
</thead>
</table>

The cost of screening colonoscopy ranged from $300 to $2,627 [87-91]. An average US price of $1,068.59 was used. There was also an additional cost of $92.06 due to adverse events [88]. Multiplying the price of each adverse event by its rate and summing over all the adverse events calculated the cost due to adverse events. The types of adverse events incorporated were perforation, serosal burn, bleed with transfusion, bleed without transfusion, and post polypectomy hemorrhage. The two costs were combined to obtain the overall cost of colonoscopy where  is the probability of receiving
a follow-up colonoscopy at age (at scheduled colonoscopies and when colonoscopies are not scheduled due to symptoms).

\[
\text{Eq. 34}
\]

Initial and surveillance [87-91] costs were included within the model for each of the three cancer stages (Local, Regional and Distant). A general terminal CRC cost was used because terminal costs are nearly the same for all three stages of CRC. A weighted average was calculated for each cancer stage based on the different costs. CRC is mostly characterized by stages but in our model we assumed local CRC to be the equivalent of stage 0 and stage 1 CRC combined, regional CRC to be equivalent to stage 2 and stage 3 CRC combined, and distant CRC to be equivalent to stage 4 CRC. From the different costs, an average initial cost was chosen for local, regional and distant CRC. The same rationale was applied for the surveillance costs of CRC from the different costs sources obtained. The costs were matched against projected costs for CRC for 2020 [92] and they fell within the ballpark of these projected costs.

3.3.6.2 Discounting

In order to best compare interventions with different costs and health effects, it is common practice to use a discounting in simulations. Discounting is a way of valuing rewards and events in the present more than the future. Both costs and QALYs were discounted. The overall outcome of the analysis was cost per QALY gained so both criteria had to be discounted. This prevents the problem of intertemporal inconsistencies, which are the inconsistencies that exist when people’s current choices affect the options they have available in the future. All costs were discounted at an annual rate of 3%. The discounting factor at each age was calculated by
where $r$ was the rate and $t$ was the age. The discounting factor was used in the colonoscopy costs to give an updated cost of

$$\text{Eq. 35}$$

Local, regional and distant discounted costs were calculated the same way. For the first year, all the costs for screening and diagnosis were incorporated and multiplied by the probability of having local cancer at that age. After screening and diagnosis the patient goes in to the first year of surveillance where surveillance costs for local CRC are added to local CRC costs and multiplied by the probability of having local cancer. This calculation continued for the surveillance years with the probability of local cancer changing for each year. For local terminal costs, the calculations were similar to the surveillance but instead we used the probability of the patient dying with local cancer at each age multiplied by the probability of local CRC. Costs were calculated similarly for regional and distant CRC. Total costs were obtained by

$$\text{Eq. 36}$$

$$\text{Eq. 37}$$

where stage=local, regional, distant, and

3.3.6.3 Quality-Adjusted Life-Years (QALYs)

A QALY is a measure of the value used for quality of life. It focuses both on quantity and quality of life given by healthcare interventions. It deals with a scaled score ranging from 0 to 1, where 1 represents a utility for perfect health and zero represents a
utility for a health state close to death. With increasing age, the utilities associated with perfect health declines because several health issues arise with increasing age. In order to calculate the QALYs, the utilities were obtained from Ness et al [88] (see Table 35).

**Table 35**: QALYs used in the cost-effective analysis.

<table>
<thead>
<tr>
<th>Patient status</th>
<th>Annual QALY Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>0.91</td>
</tr>
<tr>
<td>Dead</td>
<td>0</td>
</tr>
<tr>
<td>Local CRC</td>
<td>0.74</td>
</tr>
<tr>
<td>Regional CRC</td>
<td>0.70</td>
</tr>
<tr>
<td>Distant CRC</td>
<td>0.25</td>
</tr>
</tbody>
</table>

The disease specific health state utilities were multiplied by the probability of being in that state at each given age and the results were summed up for all the health states over all ages to obtain the total QALYs:

\[
\text{Eq. 39}
\]

where \( i = \text{healthy, dead, local, regional, distant} \)

and

\[
\text{Eq. 310}
\]

where \( i = \text{healthy, dead, local, regional, distant} \).
The QALYs were also discounted at 3% using the same discounting factor in section 3.3.6.2. The discounting factor was used in the QALYs to give updated QALYs given by

\[ \text{Eq. 39} \]

### 3.4 Cost-effectiveness Analysis

By applying the POMP model we determined the inter-colonoscopy interval that maximizes patient's QALYS per unit cost of colon-related expenses. The primary outcome of a cost–effectiveness analysis is the incremental cost- effectiveness ratio (ICER), otherwise known as the cost per QALY. This is calculated as the difference in the expected cost of two interventions, divided by the difference in the expected QALYs produced by the two interventions. \textbf{Eq. 310} gives the formula used to calculate cost per QALY gained/10,000 persons:

\[ \text{Eq. 310} \]

The outcome that yielded the most QALYs per unit cost of CRC expenses is the optimal solution. Over the years there has been an increase in articles referencing both $50,000 and $100,000 as the society’s willingness to pay (WTP) threshold for a quality adjusted life year[77, 78] so this fueled our urge to use a $50,000, $75,000 and $100,000-per-QALY threshold. In the analysis, effectiveness was the change in QALYs and the efficiency referred to the change in cost per QALY.

Once the simulation for the protocols was completed, the focus was solely on the cases where the cost per QALY gained was less than or equal to the thresholds (the cost-effective scenarios). From that pool of cost-effective scenarios, the scenario with the most QALYs gained/10,000 persons was collected. This gave us our optimal interval for the varying protocols.
Efficiency was assumed to be the change in cost over the change in QALYs.

Eq. 311

Effectiveness was assumed to be the change in QALYs gained by

Eq. 314

3.5 Sensitivity Analysis

Obtaining good impartial estimates of both costs and effectiveness of different protocols in a cost-effective analysis is very important. A sensitivity analysis is an easy approach used by modelers whose main goal is to support decision makers. Since the factors, values, and assumptions of any model can easily result in error and change; sensitivity analyses are needed to investigate these changes and errors and what effect they have on the model.

A main factor in the calculation is set to either the extreme low or extreme high and is recalculated. The simplest form of sensitivity analysis is a univariate analysis, only one main parameter is changed at one time, and correlations between parameters are not taken into account. The results obtained from the extreme values are compared to the original results. The difference in results will tell you how sensitive the results are to that particular parameter.

A one-way sensitivity analysis was performed to test the effect changes in various fundamental assumptions would have on the cost-effectiveness. In this study, changing the cost of colonoscopy, the cost of cancer, and the discount rate did the one-way sensitivity analysis. The ranges used in the sensitivity analysis were based on literature
review [87-91] (see Table 34). For the sensitivity analysis on the discount rate we tested a 0% discount rate. The entire model was analyzed using a Matlab.

CHAPTER 4

ANALYSIS OF SIMULATION PROTOCOLS

FOR COLONOSCOPY FOLLOW-UP
4.1 Introduction

Chapter 3 discussed the methods needed to analyze the cost-effectiveness of follow-up colonoscopy. The goal of this section of the research was to answer two main questions:

1. Is follow-up colonoscopy based on patient age and screening colonoscopy results cost-effective?
2. What is the cost-effective interval that maximizes patients QALYs gained by age and screening results?

4.2 Follow-up Colonoscopy based on Patient Age and Screening

Is follow-up colonoscopy based on patient age and screening colonoscopy results cost-effective?

Follow-up colonoscopy was only cost-effective after screening colonoscopy at age 50 years and at the $100,000/QALY gained threshold. Follow-up colonoscopy for screening that occurs after the age of 50 years was not cost-effective and no scenario was cost-effective at the $75,000/QALY gained or the $50,000/QALY gained thresholds.

4.2.1 Age 50

At age 50, one follow-up colonoscopy was cost effective. In Figure 41 to Figure 45 all the follow-up colonoscopies (2 to 20 years later) at age 50 for all screening findings can be seen. Looking at Figure 41 to Figure 45, only the $100,000 per QALY threshold is cost effective. At screening age 50 no cost-effective follow-up colonoscopies
were recorded for the other two thresholds. There is nothing below $80,000. Figure shows the results obtained for the cases where screening age 50 was not cost-effective.

**Figure 41:** Efficiency vs Surveillance intervals at age 50 for no neoplasia finding.
**Figure 42:** Efficiency vs Surveillance intervals at age 50 for 1 to 2 polyps only finding.

**Figure 43:** Efficiency vs Surveillance intervals at age 50 for 3 plus polyps only finding.
Figure 44: Efficiency vs Surveillance intervals at age 50 for 1 to 2 advanced polyps only finding.

Figure 45: Efficiency vs Surveillance intervals at age 50 for 3 plus advanced polyps only finding.
Based on the results, cost-effectiveness of follow-up colonoscopy at age 50 did not vary very much with the screening findings. All the follow-up colonoscopies that were cost-effective at the different screening findings occurred between 7.7 to 20 years later. A linear interpolation was done to find the exact point where the colonoscopies became cost-effective. Because the only threshold that was cost effective was $100,000, the lower and upper intervals where this threshold fell between were used to get the most cost-effective follow-up year (see **Table 42**). **Table 4-1** shows the ranges that were cost effective for each screening finding.

**Table 41**: Ranges of follow-up colonoscopies that were cost effective at age 50 for all screening findings.

<table>
<thead>
<tr>
<th>Age at screening colonoscopy (years)</th>
<th>No Neoplasia</th>
<th>1 to 2 non-advanced adenomas only</th>
<th>3+ non-advanced adenomas only</th>
<th>1 to 2 adenomas with some advanced neoplasia</th>
<th>3+ adenomas with some advanced neoplasia</th>
<th>All stages</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>7.6 – 20 years</td>
<td>7.4 – 20 years</td>
<td>7.8 – 20 years</td>
<td>8.2 – 20 years</td>
<td>8.4 – 20 years</td>
<td>7.7– 20 years</td>
</tr>
</tbody>
</table>

Follow-up colonoscopy was only cost-effective after screening colonoscopy at age 50 years at the $100,000/QALY gained threshold.

4.2.2 **Age 55**

Follow-up colonoscopy after screening at age 55, 60, 65, 70 and 75 years was not cost-effective at any of the three thresholds. Demonstrated in **Figure 46** -**Figure 410** are the scenarios for the ages 55 – 75 with the general all stages finding. These instances
were not cost effective at the 3 thresholds. The follow-up colonoscopies cost more than $100,000.

**Figure 46:** Efficiency vs Surveillance intervals at age 55 for a general all screening results- an average of all the findings combined.

**Figure 47:** Efficiency vs Surveillance intervals at age 60 for a general all screening results- an average of all the findings combined.
Figure 48: Efficiency vs Surveillance intervals at age 65 for a general all screening results- an average of all the findings combined.

Figure 49: Efficiency vs Surveillance intervals at age 70 for a general all screening results- an average of all the findings combined.
What is the cost-effective interval that maximizes patients QALYs gained by age and screening results? Two control groups were tested to answer this question.

4.3.1 Results for Control Group 1

From the pool of colonoscopies that were cost-effective (see section 3.3.5.1.1), the one that produced the most QALYs was considered to be optimal. A quadratic interpolation was done to determine what that interval was that yielded the most QALYS.
Three points were chosen - the maximum point and the two points surrounding the maximum - and were interpolated to obtain the interval. The interpolation was done in Matlab.

4.3.2 Age 50

The cost-effective range for the $100,000 threshold for age 50 with no neoplasia was 7.6 to 20 years. Out of that range, what interval gave the highest QALYs (optimal solution)? Figure 411 shows the QALYs-gained per 10,000 patients for the follow-up years. In the figure the three points that were used for the interpolation can be seen highlighted in the box. The maximum falls somewhere between 6 and 10 years (see Table 42). For all other screening findings, (see to Figure 412 to Figure 415), the same procedure was used to obtain the optimal interval.

Figure 411: Effectiveness vs Surveillance intervals (ΔQALY) at age 50 with no neoplasia.
Figure 412: Effectiveness vs Surveillance (ΔQALY) at age 50 with 1 to 2 small polyps only.

Figure 413: Effectiveness vs Surveillance (ΔQALY) at age 50 with 3 plus small polyps only.
Figure 414: Effectiveness vs Surveillance (ΔQALY) at age 50 with 1 to 2 advanced polyps only.

Figure 415: Effectiveness vs Surveillance (ΔQALY) at age 50 with 3 plus advanced polyps only.
The inter-colonoscopy intervals were largely insensitive to the screening colonoscopy results; inter-colonoscopy intervals stratified by the screening result were all within a 0.5-year range (see Table 42). Because of this insensitivity and for simplicity, the average follow-up interval is reported and the intervals stratified by screening result reported in Table 42.

**Table 42:** Inter-colonoscopy intervals that maximize QALYS gained for the $100,000 threshold.

<table>
<thead>
<tr>
<th>Age at screening colonoscopy</th>
<th>No Neoplasia</th>
<th>1 to 2 non-advanced adenomas only</th>
<th>3+ non-advanced adenomas only</th>
<th>1 to 2 adenomas with some advanced neoplasia</th>
<th>3+ adenomas with some advanced neoplasia</th>
<th>All stages</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 years</td>
<td>8.5 years (83.98)</td>
<td>8.5 years (84.04)</td>
<td>9.0 years (83.29)</td>
<td>8.4 years (81.25)</td>
<td>8.9 years (80.71)</td>
<td>8.5 years (83.96)</td>
</tr>
</tbody>
</table>

Following screening at age 50 years the most QALYs gained at the $100,000/QALY threshold was at one follow-up colonoscopy with interval of 8.5 years. There were 83.96 QALYs gained per 10,000 people at that threshold. There were no cost-effective follow-up colonoscopies for screening at older ages. **Table 42** shows the interval where the most QALYS were gained for each finding.

4.3.3 **Age of final colonoscopy**
There was never a cost-effective follow-up colonoscopy after age 60 years. At the $100,000/QALY threshold, after screening at age 50 for all the stages finding, no follow-up colonoscopies were recorded after age 58.5 years.

### 4.4 Sensitivity Analysis

From the sensitivity analysis the model was relatively insensitive to high screening colonoscopy costs and low cancer cost. When the higher end of the colonoscopy cost was used, nothing was cost-effective at any screening age at any of the three thresholds. When using the low end of the cost of cancer, at screening age 50 years, the cost-effective window was recorded between 10 to 20 years at the $100,000 threshold, compared to 7.7 to 20 years from the original analysis. Also the recommended follow-up interval dropped from 8.5 years to 8.3 years, and the QALYS gained/10,000 persons decreased by 0.1. Nothing became cost-effective at the $75,000 or $50,000 threshold.

The model was however sensitive to lower screening colonoscopy costs, higher cancer cost and the 0% discount rate (see Table 43). In the table, a single number “x” represents one follow-up colonoscopy was cost-effective, two numbers “x, y” represents two follow-up colonoscopies were cost-effective and so on.

**Table 43:** Inter-colonoscopy intervals for sensitivity analyses.

<table>
<thead>
<tr>
<th>Screen Age</th>
<th>Outcome</th>
<th>Colonoscopy Cost</th>
<th>Cancer Cost</th>
<th>Discount Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Age</td>
<td>Recommended follow-up interval</td>
<td>Range of cost-effective follow-up colonoscopy</td>
<td>Gain in QALYs / 10,000 persons</td>
<td>Gain in QALYs / 10,000 persons</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------</td>
<td>---------------------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>50 years</td>
<td>8.5 years</td>
<td>2 – 20 years</td>
<td>Not cost-effective</td>
<td>84.0</td>
</tr>
<tr>
<td>55 years</td>
<td>6.8</td>
<td>2 – 20 years</td>
<td>Not cost-effective</td>
<td>75.1</td>
</tr>
<tr>
<td>60 years</td>
<td>5.6 years</td>
<td>2 – 20 years</td>
<td>Not cost-effective</td>
<td>62.6</td>
</tr>
<tr>
<td>65 years</td>
<td>4.3 years</td>
<td>2 – 14 years</td>
<td>Not cost-effective</td>
<td>47.6</td>
</tr>
</tbody>
</table>

$75,000 / QALYs gained

<table>
<thead>
<tr>
<th>Age</th>
<th>Recommended follow-up interval</th>
<th>Range of cost-effective follow-up colonoscopy</th>
<th>Gain in QALYs / 10,000 persons</th>
<th>Gain in QALYs / 10,000 persons</th>
<th>Gain in QALYs / 10,000 persons</th>
<th>Gain in QALYs / 10,000 persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 years</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
</tr>
<tr>
<td>Age</td>
<td>55 years</td>
<td>50 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>Effective</td>
<td>Effective</td>
<td>Effective</td>
<td>Effective</td>
<td>(307.30)</td>
<td></td>
</tr>
<tr>
<td>$50,000 / QALYs gained</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>8.2 years (157.14)</td>
<td></td>
</tr>
<tr>
<td>$50,000 / QALYs gained</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>Not cost-effective</td>
<td>10.3 years (194.12)</td>
<td></td>
</tr>
</tbody>
</table>

When the low end of the colonoscopy cost was used, more screening ages (55, 60, and 65 years) were found to be cost-effective at the $100,000 threshold and it was also observed the interval for screening in the age 50 group widened from 7.7 to 20 years to 2 to 20 years. With the screening ages that were cost-effective, the recommended follow-up intervals and number of QALYS gained/10,000 persons decreased with increasing age. With the high end of cancer cost, screening age 55 became cost effective with follow-ups 10 to 14 years later with an average 5.8 follow-up interval and 75.6 QALYS gained/10,000 persons. For screening age 50, the range closed by a mere 0.3 years with the same follow-up interval and QALYs gained from the original analysis. With the 0% discount rate showed as many as 4 follow-up colonoscopies were cost-effective at screening age 50 at the $100,000 threshold with 437.22 QALYs gained/10,000 persons. More screening ages also became cost-effective at the $50,000 and $75,000 thresholds. Based on the sensitivity analysis we can see that with increasing threshold on the societal willingness to pay for a QALY, more follow-up colonoscopies are possible with higher QALYs gained/10,000 persons.

4.4.1 Low Screening Colonoscopy Costs
After the cost of screening colonoscopy was adjusted to its ultimate low, it was observed that more screening ages became cost effective for the $100,000 threshold. From the cost of colonoscopy used in the original simulation, only age 50 was cost-effective at the $100,000 threshold. After the parameters changed, ages 50, 55, 60, and 65 became cost effective at the $75,000 and $50,000 threshold as well. With the cost of colonoscopy at its lowest we can see that follow-up colonoscopies at screening ages 50, 55, 60, and 65 being cost-effective at all the thresholds.

**Figure 416** to **Figure 419** show the surveillance intervals vs. efficiency for screening at age 50, 55, 60, and 65, respectively, for the lowest screening colonoscopy cost. Below the circled area on the figures shows the region where the follow up intervals that was cost effective at the $75,000 threshold and the line shows the region where the follow up intervals that was cost effective at the $50,000 threshold.

![Graph](image)  

**Figure 416:** Efficiency vs Surveillance intervals at age 50 with all screening results.
Figure 417: Efficiency vs Surveillance at age 55 with all screening results.

Figure 418: Efficiency vs Surveillance at age 60 with all screening results.
Figure 419: Efficiency vs Surveillance at age 65 with all screening results.

Older screening ages become cost effective with the lowering of the cost of colonoscopy. Also the range of cost-effectiveness widens with lower screening colonoscopy cost (as seen in Table 43). At screening age 65, there was a slightly smaller cost-effective range for the $100,000 and $75,000 thresholds. However, the inter-colonoscopy intervals are sensitive to age. As the screening ages increased the inter-colonoscopy interval decreases.

4.4.2 High Cancer Cost

With the high end of cancer cost, screening at age 55 was cost effective 10 to 14 years later with an average 5.8 follow-up intervals and 75.6 QALYs gained/10,000 persons. For screening age 50, the range from the original analysis (see Table 41)
decreased by a mere 0.3 years with the same follow-up interval and QALYs gained from the original analysis (see Table 43).

4.4.3 Zero Percent Discount Rate

From Table 43, the 0% discount rate showed as many as 4 follow-up colonoscopies were cost-effective at screening age 50 at the $100,000 threshold. A total of 437.22 QALYs were gained/10,000 persons. More screening ages also became cost-effective at the $50,000 and $75,000 thresholds.

The optimal number of follow-up colonoscopies and their timing varied based on the cost-effectiveness threshold with a trend of more colonoscopies and more frequently as the cost-effectiveness threshold increased.

The optimal number of colonoscopies following a screening colonoscopy at age 50, increased from a single follow-up colonoscopy at the $50,000/QALYs gained threshold to two follow-up colonoscopies at the $75,000/QALYs gained threshold to four follow-up colonoscopies at the $100,000/QALYs threshold. At the $50,000/QALYs gained threshold, the optimal inter-colonoscopy interval increased from 8.5 years to 10.3 years with more than double increase in QALYs gained/10,000 persons.

At the $75,000/QALYs gained threshold, we moved from nothing being cost-effective for that threshold with the 3% discount rate to 2 follow-up colonoscopies being cost-effective yielding over three times as many QALYs gained/10,000 persons compared to the $50,000 threshold. An additional screening age of 55 also becomes cost-effective at that threshold.

At the $100,000/QALYs gained threshold, we moved from nothing being cost-effective for that threshold with the 3% discount rate to 4 follow-up colonoscopies and
also additional screening ages being cost-effective yielding over four times as many QALYs gained/10,000 persons compared to the $50,000 threshold.
CHAPTER 5

COST-EFFECTIVENESS OF FECAL IMMUNOCHEMICAL TESTS (FITS)

5.1 Introduction

In section 2.2.2, non-endoscopic tests for CRC were discussed. Some acclaimed cost-effective policies for CRC in the US include colonoscopy, sigmoidoscopy and stool tests [62]. Stool based screening tests are cost effective, non-invasive and have good accessibility [48]. Fecal tests look for hidden blood in the stool. This could indicate an early sign of CRC. Since fecal tests have poor sensitivity to premalignant lesions, they must be repeated more often usually 1-2 years [48]. Considering the shorter intervals of fecal testing required due to poor sensitivity, but enabled by the low cost of the tests; coupling this tests with colonoscopy is expected to be more cost-effective and eventually cost saving.

Fecal immunochemical testing (FIT) is a great alternative and complement to colonoscopy according to the US Multi-Society Task Force [93]. It is a non-invasive test. When used as a single application test, FIT is about 79% sensitive for detecting cancer and about 20% - 30% sensitive for detecting advanced neoplasia [93].

For the cost-effectiveness analysis for FIT we focused on two main goals with the same concept as the colonoscopy analysis discussed in Chapter 4 (see Figure 51). Since
FITs are different from colonoscopies and do not have a follow-up regimen, our main goals were aimed around:

1. After a screening colonoscopy, when should FITs be administered, and
2. What is the best time to screen for FITs?

**Figure 51:** Overview of FIT methods.

### 5.2 Overview of simulation process for FIT

The colonic neoplasia growth model and mortality rate model in sections 3.3.1 and 3.3.2 were the same for the FIT simulation process.

#### 5.2.1 POMP Model for FIT

Similar to section 3.3.3, leading to FIT screening, the adenoma growth and mortality rate models are applied to the patients. A patient can die from natural causes, remain in a healthy state, develop symptoms and develop polyps or CRC. These patients are not placed in the patient pool initially, because each hypothetical patient is considered healthy until they receive a screening FIT. A FIT test can either be positive or negative. In our model, a positive test for a FIT triggered a colonoscopy and the patient had an
interval before a next FIT was done. Similarly if a negative FIT were recorded, the patient would have an interval until the next FIT. Since a colonoscopy is done after a positive FIT, the intervals between FITs is greater when a FIT is positive than when it is negative because the colonoscopy would have possibly removed most of what was there; leaving a longer window before being tested again.

5.2.2 FIT Protocol

For the analysis we tested different start ages, stop ages and also different intervals for if the test was positive or negative. The start age to receive a FIT was set to be 50, 55, 60, 65, 70 or 75 years. The positive step, which is when to get another FIT if results were positive, was set at 4, 5 or 6 years for each screening age and the negative step, which is when to get another FIT if results are negative, was set at 2, 3 or 4 years for each screening age. The age to stop FITs was set at 70, 75, 80, 85, or 90 years. The cost per QALYs gained as a function of start age, stop age, positive interval, and negative interval were identified. The probabilities collected from the POMP model were incorporated into the costs and QALYs results to perform the cost-effectiveness analysis.

5.2.3 Costs

Since, a positive FIT triggers a colonoscopy; colonoscopy related costs (see Table 34) were incorporated in the FIT analysis. The cost of a FIT kit ranged from $10- $50 [94-97]. An average cost of $19.50 was used in the analysis. The cost of FIT was obtained from

Eq. 51
where \( P \) is the probability of receiving a FIT at age \( a \) (for scheduled FITs and when FITs are not scheduled but given due to symptoms).

The costs for FIT were added to the calculations for the total costs in section 3.3.6.1. Total costs were then obtained by

\[
\text{Eq. 52}
\]

The FIT costs were discounted at an annual rate of 3\% and calculations were obtained in the exact manner as section 3.3.6.1 with

\[
\text{Eq. 53}
\]

5.2.4 QALYs

Section 3.3.6.3 explains the process that was used in calculating the QALYs.

5.2.5 Cost-effectiveness Analysis

To perform the cost-effective analysis for screening using FIT only, we run the model using the control group that was not screened. This analysis aided in answering the question of the best time for screening with just FITs.

5.2.5.1 Control Group: No screening from age 0 until death or age 100

In this case, our control group consists of patients who had no screening done. The adenoma growth model, mortality rate model, and symptoms calculations were applied to patients from age 0 to death or 100. The costs and QALYs were obtained for patients who received FITs versus those who just had symptoms, died naturally or died from CRC.
By applying the POMP model, we expected to find the most favorable start age, positive step, negative step, and stop age for FIT screening that would maximize patient’s QALYs per unit cost of colon-related expenses. **Eq. 54** gives the formula used to calculate cost per QALY gained/10,000 persons.

**Eq. 54**

For this control group, considering start age, stop age, positive interval, and negative interval, we examined the outcome that yielded the most QALYs per unit cost of CRC expenses. The same thresholds discussed in section 3.4 were used as the societal willingness to pay (WTP) for a QALY. In the analysis, effectiveness was the change in QALYs and the efficiency referred to the change in cost per QALY.

## 5.2.6 Results from FIT only Protocol

**Table 51:** FIT only Analysis.

<table>
<thead>
<tr>
<th>Recommended start age</th>
<th>When to return if FIT is positive</th>
<th>When to return if FIT is negative</th>
<th>Recommended stop age</th>
<th>QALYs</th>
</tr>
</thead>
<tbody>
<tr>
<td>$75,000 / QALYs gained threshold</td>
<td>55 years</td>
<td>6 years</td>
<td>2 years</td>
<td>85 years</td>
</tr>
</tbody>
</table>
With the results from this analysis, since we are trying to obtain the best protocol for FIT only, we can only evaluate analysis on a single start age, negative step, positive step, and stop age at a particular time. In order to do an analysis comparing start ages, we calculated QALYs that were not discounted to obtain the actual QALYs gained.

There was no FIT only protocol that was cost-effective at the $50,000 threshold. However at the $75,000 and $100,000 thresholds there were some protocols that were cost-effective. From the cost-effective pool, the protocol that yielded the most QALYs for the threshold is the best protocol for FIT only. A Matlab function was written to obtain the protocol with the highest QALYs, from there the two surrounding points to the maximum was used to show the maximum protocol. Figure 5-2 shows the QALYs gained for each parameter tested for the best protocol for the $75,000 threshold. In Figure 5-3 shows the cost-effectiveness at the $75,000 threshold. The same technique was done for the $100,000 threshold.

From the results obtained in Table 5-1, with increasing thresholds there is an increase in start age and a decrease in frequency to return if the FIT is positive.
Figure 52: QALYs gained for the best protocol for FIT only.

Figure 53: Efficiency Vs Age for the best FIT only protocol.

5.2.7 Combination of Colonoscopy and FIT Tests
In the beginning of this chapter, I discussed how coupling the two screening tests could be more cost-effective due to the cheap price of FIT and its high sensitivity. We performed a cost-effective analysis on a combination of the two tests to determine what effect coupling the tests would have. The colonoscopy and FIT combination involved patients receiving a colonoscopy at the start ages assigned and only FITs thereafter. The methods for the FITs and the colonoscopy remained the same as discussed earlier.

With this combination analysis, we had two control groups in which we analyzed to determine the effectiveness:

1. Control group 1: Screening colonoscopy only at the start age, and

2. Control group 2: No screening from age 0 until death or age 100.

With the control group where a screening colonoscopy was done at the start age, every single case was cost effective at the $50,000 threshold making it cost-effective at $75,000 and $100,000 as well. At each screening age, Table 52 shows the recommended time to get a FIT, the frequency and when to stop. We can see FIT intervals become shorter with growing age and an increase in QALYs. We also observed that when a FIT is negative, it is recommended to come back every 2 years consistently for all age groups which justify current recommendations[48]. If a FIT comes back negative it may have missed something so we expect the intervals to be shorter.

Table 52: Combination of FIT and colonoscopy analysis with control group 1.
With the control group where no screening was done from age 0 to death or age 100, every single case was cost effective at the $50,000 threshold making it cost-effective at $75,000 and $100,000 as well. The protocol that yielded the most QALYs gained/10,000 persons (69.23) is seen in Table 53.

**Table 53:** Combination of FIT and colonoscopy analysis with control group 2.
CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

We obtained the optimal inter-colonoscopy interval at the $100,000 threshold following a screening colonoscopy at age 50 years. The interval is approximately 8.5 years. What that means is that after a screening colonoscopy at age 50 a follow-up is recommended at age 58.5. Clinically for advanced findings (3 or more adenomas found at screening colonoscopy), the recommended interval to return is short (~3 years) [98, 99]. However, with this study, the follow-up interval observed was consistent across screening colonoscopy results for an average risk patient [62]. The shortening of clinical recommendations is based on data that advanced neoplasia are more prevalent following screening colonoscopy with advanced neoplasia [71, 98, 100]. The mathematical model also predicts a higher risk of CRC in patients with advanced neoplasia but, in the simulations, most of the patients who developed cancer got a follow-up colonoscopy due to symptoms prior to the recommended interval. Because of the inclusion of colonoscopy due to symptoms, the systematic follow-up of all patients at an earlier time is not necessarily cost-effective on the entire population level.

After screening age 50, follow-up colonoscopies are not cost effective. However, with a 0% discount rate we saw as many as four follow up colonoscopies being cost effective at the $100,000/QALY threshold. The last follow-up colonoscopy from
screening at age 50 was recorded at 68.8 years. This finding can be validated because mortality rate increases with age so there is not much time to see preventative effects because patients get old and die eventually. Studies have shown that the benefits of colonoscopies decrease with age [66, 101]. The US Preventive Task Force (USPTF) recommends screening to end at 75 years [66]. The average life expectancy in the US is approximately 79 years so the impact of CRC prevention and early detection due to colonoscopy will likely be minimal after age 75 years. In addition, after age 75 years people tend to have more health issues and undergoing colonoscopy may not be the most comfortable experience plus the rewards would not be significant as discussed earlier. From the sensitivity analysis (see Table 43) we also saw that with lower screening colonoscopy costs and higher cancer costs we get follow up colonoscopies at later ages.

Not surprisingly, the number of follow-up colonoscopies will increase if one is willing to pay more for a unit of health. From increased amounts in what one is willing to spend, we can see that more colonoscopies can occur leading to shorter inter-colonoscopy intervals. For decades $50,000 has been used as a benchmark for the cost of a unit of health [77-79]. Using the benchmark of $50,000 per QALY gained may be somewhat outdated due to inflation and economic growth [77]. It is a reasonable enough cost but based on our study that makes the necessary treatment not cost-effective. By spending more you can receive more benefit. It should be noted that some people would spend more than $100,000 to ensure better health. The threshold used, is solely based on people’s values and how they approach risk.

More benefit would mean more QALYs gained from more colonoscopies. The more tests one can get the more information that can be gathered to aide in preventing or
reducing CRC incidence [41, 64]. Cost and benefit are directly proportional as shown in the study. More people are getting screening colonoscopy and will have a lifetime of potential follow-up colonoscopy. Early follow-up colonoscopy in patients who are high risk is extremely crucial[102].

Our results indicates that people should receive a follow-up colonoscopies a little sooner than 10 years and we did obtain data quantifying those intervals based on age at screening and also based on the findings at that screening. With advancing technology and medical improvements, spending a little more for a unit of life if you can is not a bad idea. Based on the study continuing screening after the age of 75 years is not recommended. Follow-up colonoscopies may give more years to live given the appropriate protocol based on age and findings from screening.

Combining FITs and colonoscopy turned out to be very cost-effective. Each of the screening ages tested was cost effective with the combination producing more QALYs with increasing age. From the obtained results, coupling the two screening strategies will be the best cost-effective outcome.

The current study on follow-up colonoscopy is the first to quantitatively analyze the frequency of follow-up colonoscopies based on age and findings from screening results.

6.2 Future Work

With this research we focused on rate of development of CRC in a general aspect. In the future, the focus on the development of the disease will be specified by location in the colon (distal or proximal colon). Critical data will have to be obtained on cancer in the distal vs proximal colon. A recent study shows that younger patients get more
cancers in the distal colon but they may take longer to develop, while older patients are more prone to proximal cancers [103]. This study also showed trends with an increase in diagnosis of CRC within adult less than 55 years. This can open opportunities for investigating the younger birth cohorts primarily focusing on site of the disease. The rate at which CRCs develop in the distal or proximal colon may influence the use of screening techniques for the disease.

Also, since the study is limited to Veteran Affairs (VA) data, which is predominantly white males, a next direction would be to look into databases with populations that are more inclusive of gender and other races and apply the model to see what the effect would be on these populations. We can test the model on different populations to compare results from different populations. If we can obtain data to include a population of different races and different sexes, it may make a stronger argument as to recommendations for an entire population as opposed to part of a population. Obtaining adequate diverse data from the VA may pose a bit of difficulty but it may help understand greatly the other populations not covered by the VA data.

6.3 Limitations of the study

A limitation of the study is that the results rely on an adenoma growth model developed using data from a single Veterans Affairs (VA) hospital. While this model is accurate for this population, the population consisted of few female and non-white patients. Outside of the VA, the results obtained may not be applied in a broad aspect but the methodology and model basis can be applied to data sets with similar characteristics.
APPENDIX A

MATLAB CODE USED FOR MODELING

A.1 Overview

This entire research focused on cost-effectiveness of two different screening techniques for colorectal cancer and a combination of the two. We used Matlab to simulate colonoscopies and FITs in hypothetical patients. The code required to run the model and the protocols tested will be presented here in the Appendix. A total of 5 main scenarios were analysed for the cost-effectiveness of this project:

1. Given a screening colonoscopy when to have follow-up colonoscopies,
2. Best time for colonoscopy screening,
3. Best time for FIT screening,
4. Given a screening colonoscopy when to have follow-up FITs,
5. Best time for screening colonoscopy and FIT.

Each scenario had a main function and subroutines to run that main function. The subroutines varied for each scenario. However, some subroutines were constant for all scenarios. I will go over the code needed to run the combination of colonoscopy and FIT protocols. The code can be tailored to do colonoscopy only or FIT only.
A.1.1 Given a screening colonoscopy when do you have follow-up FITs

For the main function for the combination of screening tests, before cost and/or QALYs are calculated, the following are subroutines that will need to run.

This is the main function that gives the cost and QALYs output. Withinin this main function there a several smaller functions, which will be discussed below. Below is a list of the functions that are needed to run the main function:

1. STATE_DEFINITIONS
2. K_MARKOV_MOD
3. P_F_A
4. COLONOSCOPY
5. STATE_GROUP
6. P_HEALTHY_DEATH
7. P_HEALTHY_CRC
8. P_CRC_DEATH
9. STATE_DEFINITIONS
10. STATE_TRANSITIONS
11. TRANSITION_RATES
12. TREATMENT_DEFINITIONS
13. TREATMENT_GROUP
14. P_PATIENT_COLON_FIT_COMBO
15. TOTAL_QALYS_DISCOUNTING_FIT
16. TOTAL_COSTS_FIT_DISCOUNTING_RATE

1. STATE_DEFINITIONS:
This function gives each combination of adenomas and CRC to determine the possible state of the colon. This function gives a total 924 states.

2. **K_MARKOV_MOD:**

This function identifies the transition rates for the series of transition from diminutive adenoma (<5 mm) to medium adenoma (6-9 mm) to large/advanced adenoma (>10 mm) to CRC. These rates are used to obtain the colonic state probability vector.

3. **P_F_A**

This function calculates the probability of finding an actual state during a colonoscopy.

4. **COLONOSCOPY_6**

The colonoscopy function goes through each state and sees when a colonoscopy should be performed. It also removes any abnormalities found during a colonoscopy.

5. **STATE_GROUP**

The state group function regroups the 924 states from **STATE_DEFINITIONS** into the 8 findings categories: (1) no adenomas, (2) 1-2 non-advance adenomas only, (3) 3+ non-advanced adenomas only, (4) 1-2 adenomas with some large/advanced adenomas, (5) 3+ adenomas with some large/advanced adenomas, (6) local CRC, (7) regional CRC, and, (8) distant CRC.

6. **P_HEALTHY_DEATH**

In this function the probability of a healthy patient dying from non CRC reasons is calculated.

7. **P_CRC_DEATH**
In this function the probability of dying from the disease is given.

8. P_HEALTHY_CRC

In this function the probability of a healthy patient getting CRC is calculated. This function calculates the colon state as a person ages, adjusts the colon state when there is a colonoscopy and calculates the probability of developing cancer.

9. STATE_DEFINITIONS

This function gives all the combination of states.

10. TRANSITION_RATES

In this function the rate at which adenoma forms and progresses is calculated. Symptom development rates are also developed in this function.

11. TREATMENT_DEFINITIONS

This function gives the combination of positive and negative FITs. With FITs order matters. There are a total of 127 FIT treatments.

12. P_Patient_COLON_FIT_COMBO

This function calculates the probabilities from age 0 to death or age 100. This function has a series of blocks computing different probabilities. It calculates healthy mortality for the year, cancer related mortality, colon state as a person ages, and the probability of colonoscopy based on symptom development. In this function is where patients get a screening colonoscopy then followed by a FIT follow-up. It also calculates the yearly probabilities at each block.

13. TOTAL_COSTS_FIT_DISCOUNTING

This function calculates the total costs at each time step.
14. TOTAL_QALYs_DISCOUNTING_FIT

This function calculates the total QALYs at each time step.

15. FIT_MATRIX ()

This function creates a matrix to know when a positive interval and a negative interval for a FIT are taken for each treatment.

16. FIT_AGE (start_age, positive_step, negative_step)

This function gives you all the combinations for the age of each positive and negative interval.
MAIN FUNCTION

global n_states max_state state_list states

n_states = 6;

[states, state_list, max_state] = STATE_DEFINITIONS_6()

global K_markov

[K_markov] = K_markov_mod();

global dt age_max total_ages age_min

age_min=0;

age_max=100;

dt=1;

total_ages= (age_max-age_min)/dt +1;

global Initial(Local) CRC_costs Surveillance_Local_CRC_costs

Initial_Regional_CRC_costs Surveillance_Regional_CRC_costs

Initial_Distant_CRC_costs Surveillance_Distant_CRC_costs

Terminal_CRC_costs Cost_of_colonoscopy Cost_Adverse_effects Cost_of_FIT

Initial_Local_CRC_costs=20247.2;

Surveillance_Local_CRC_costs=1305.04;

Initial_Regional_CRC_costs=26007.5;

Surveillance_Regional_CRC_costs=2346.72;

Initial_Distant_CRC_costs=30085.2;

Surveillance_Distant_CRC_costs=15057;

Terminal_CRC_costs=23002.35;

Cost_of_colonoscopy=1068.59;
Cost_Adverse_effects = 141.79;
Cost_of_FIT = 19.5;

global QALYs_healthy QALYs_dead QALYs_Local_CRC QALYs_Regional_CRC QALYs_Distant_CRC
QALYs_healthy = 0.91;
QALYs_dead = 0;
QALYs_Local_CRC = 0.74;
QALYs_Regional_CRC = 0.70;
QALYs_Distant_CRC = 0.25;

global p_f_a colonoscopy state_group
[p_f_a] = P_F_A_6();
[colonoscopy] = COLONOSCOPY_6();
[state_group] = STATE_GROUP_6();

global P_local_death P_regional_death P_distant_death
[P_local_death, P_regional_death, P_distant_death] = P_CRC_DEATH;

global P_healthy_death
P_healthy_death = P_HEALTHY_DEATH();

Cost_matrix = zeros(6,1);
QALY_matrix = zeros(6,1);

Cost_matrix_ref = zeros(6,1);
QALY_matrix_ref = zeros(6,1);

Actual_QALYs_matrix = zeros(6,1);
Delta_Cost_Qaly_matrix = zeros(6,1);
i = 0;

for start_age=50:5:75
    i = i + 1;

j = 0;

for positive_step=4:6:10
    j = j + 1;

k = 0;

for negative_step=2:4:10
    k = k + 1;

l = 0;

for stop_age=70:5:90
    l = l+1;

%%%%%%%%Do nothing reference case%%%%%%%%

[P_death, P_Asymptomatic, P_local, P_regional, P_distant, PColonoscopy] = P_Patient_FIT_ref_max_1();

total_costs_ref = TOTAL_COSTS_FIT_discounting(P_local, P_regional, P_distant, PColonoscopy, zeros(101,1), start_age);

total_qalys_ref = TOTAL_QALYs_discounting_FIT (P_local, P_regional, P_distant, P_death, P_Asymptomatic, start_age);

Cost_Qalys_ref = total_costs_ref / total_qalys_ref;

Cost_matrix_ref = total_costs_ref;

QALY_matrix_ref = total_qalys_ref;
P\_death\_ref = P\_death;
P\_Asymptomatic\_ref = P\_Asymptomatic;
P\_local\_ref = P\_local;
P\_regional\_ref = P\_regional;
P\_distant\_ref = P\_distant;
P\_colonoscopy\_ref = P\_colonoscopy;

%%%%%%Colonoscopy then FIT scenario%%%%%%

[P\_death, P\_Asymptomatic, P\_local, P\_regional, P\_distant,
P\_colonoscopy,P\_FIT ] = P\_Patient\_COLON\_FIT\_combo(start\_age, positive\_step,
negative\_step,stop\_age);

total\_costs = TOTAL\_COSTS\_FIT\_discounting(P\_local, P\_regional,
P\_distant, P\_colonoscopy,P\_FIT,start\_age);

total\_qalys = TOTAL\_QALYs\_discounting\_FIT (P\_local, P\_regional,
P\_distant, P\_death, P\_Asymptomatic,start\_age);

Actual\_QALYS = TOTAL\_QALYs(P\_local, P\_regional, P\_distant,
P\_death, P\_Asymptomatic);

Cost\_Qalys = total\_costs / total\_qalys;

Cost\_matrix(i,j,k,l) = total\_costs;

QALY\_matrix(i,j,k,l) = total\_qalys;

Actual\_QALYs\_matrix (i,j,k,l) = Actual\_QALYS;

Toc
\[
\text{Delta\_Cost\_Qaly\_matrix}(i,j,k,l) = \frac{(\text{Cost\_matrix}(i,j,k,l) - \text{Cost\_matrix\_ref})}{(\text{QALY\_matrix}(i,j,k,l) - \text{QALY\_matrix\_ref})};
\]

\text{toc}

end

end

end

end
Here I will go into the smaller functions that are needed within the main function.

17. STATE_DEFINITIONS()

This function gives each combination of adenomas and CRC to determine the possible state of the colon. This function gives a total 924 states.

function [states, state_list, max_state] = STATE_DEFINITIONS()

global n_states

states =
zeros(n_states+1,n_states+1,n_states+1,n_states+1,n_states+1,n_states+1);
state = 0;
for i1=0:n_states
    for i2=0:n_states
        for i3=0:n_states
            for i4=0:n_states
                for i5=0:n_states
                    for i6=0:n_states
                        if (i1>=i2 && i2>=i3 && i3>=i4 && i4>=i5 && i5>=i6)
                            state = state + 1;
                            states(i1+1,i2+1,i3+1,i4+1,i5+1,i6+1) = state;
                            state_list(state,:) = [i1 i2 i3 i4 i5 i6];
                        end
                    end
                end
            end
        end
    end
end
max_state = state;
18. K_markov_mod()

This function identifies the transition rates for the series of transition from diminutive adenoma (<5 mm) to medium adenoma (6-9 mm) to large/advanced adenoma (>10 mm) to CRC. These rates are used to obtain the colonic state probability vector.

function [K_markov_new] = K_markov_mod()

global max_state

global state_list

global states

load K_markov

K_markov_new = zeros (max_state, max_state,101);

for start_state=1:max_state
    count_4_start = 0;
    count_5_start = 0;
    count_6_start = 0;

    for i=1:6
        if state_list(start_state,i+1)==4
            count_4_start = count_4_start + 1;
        end
        if state_list(start_state,i+1)==5
            count_5_start = count_5_start + 1;
        end
    end

end
if state_list(start_state,i+1)==6
count_6_start = count_6_start + 1;
end
end

for finish_state=1:max_state
    count_4_finish = 0;
    count_5_finish = 0;
    count_6_finish = 0;
    for i=1:6
        if state_list(finish_state,i+1)==4
            count_4_finish = count_4_finish + 1;
        end
        if state_list(finish_state,i+1)==5
            count_5_finish = count_5_finish + 1;
        end
        if state_list(finish_state,i+1)==6
            count_6_finish = count_6_finish + 1;
        end
    end
    equiv_start_1 = 1 + min(4,state_list(start_state,2));
equiv_start_2 = 1 + min(4,state_list(start_state,3));
equiv_start_3 = 1 + min(4,state_list(start_state,4));
equiv_start_4 = 1 + min(4,state_list(start_state,5));
equiv_start_5 = 1 + min(4, state_list(start_state, 6));
equiv_start_6 = 1 + min(4, state_list(start_state, 7));
equiv_start =
states(equiv_start_1, equiv_start_2, equiv_start_3, equiv_start_4, equiv_start_5, equiv_start_6);
if count_5_start==count_5_finish && count_6_start==count_6_finish
  equiv_finish_1 = 1 + min(4, state_list(finish_state, 2));
equiv_finish_2 = 1 + min(4, state_list(finish_state, 3));
equiv_finish_3 = 1 + min(4, state_list(finish_state, 4));
equiv_finish_4 = 1 + min(4, state_list(finish_state, 5));
equiv_finish_5 = 1 + min(4, state_list(finish_state, 6));
equiv_finish_6 = 1 + min(4, state_list(finish_state, 7));
equiv_finish =
states(equiv_finish_1, equiv_finish_2, equiv_finish_3, equiv_finish_4, equiv_finish_5, equiv_finish_6);
for age=1:101
  K_markov_new(finish_state, start_state, age) = K_markov(equiv_finish, equiv_start, age);
end
else
  diff = zeros(1, 6);
diff(1) = abs(state_list(finish_state, 2) - state_list(start_state, 2));
diff(2) = abs(state_list(finish_state, 3) - state_list(start_state, 3));
diff(3) = abs(state_list(finish_state, 4) - state_list(start_state, 4));
diff(4) = abs(state_list(finish_state, 5) - state_list(start_state, 5));
diff(5) = abs(state_list(finish_state, 6) - state_list(start_state, 6));
diff(6) = abs(state_list(finish_state,7)-state_list(start_state,7));

if (sum(diff))==1 && finish_state>start_state
    for i=1:6
        if diff(i)==1
            change = state_list(start_state,i+1);
            if change==4
                P = 0.22;
            elseif change==5
                P = 0.50;
            end
        end
    end
end

for age=1:101
    dP = P * K_markov_new(start_state,start_state,age);
    K_markov_new(start_state,start_state,age) =
        K_markov_new(start_state,start_state,age) - dP;
    K_markov_new(finish_state,start_state,age) =
        K_markov_new(finish_state,start_state,age) + dP;
end
end
end
end
19. P_F_A()

This function calculates the probability of finding an actual state during a colonoscopy.

function [p_f_a] = P_F_A()

global max_state

global state_list

global states

n_trajectories = 6;

sen_s = 1.4627;

sen_m = 1.8833;

sen_l = 3.1216;

sen_c = 3.1216;

sen_a_f = -1.9947;

sen_p = -2.8670;

n_preps = 2;

p_f_a = zeros(max_state,max_state,n_preps);

for prep = 1:2
if prep == 1
    sen_prep = sen_a_f;
else
    sen_prep = sen_p;
end
for i=1:max_state
    A1 = state_list(i,2);
    A2 = state_list(i,3);
    A3 = state_list(i,4);
    A4 = state_list(i,5);
    A5 = state_list(i,6);
    A6 = state_list(i,7);
    for i1=0:1
        if i1==1
            f1 = state_list(i,2);
        else
            f1 = 0;
        end
        for i2=0:1
            if i2==1
                f2 = state_list(i,3);
            else
                f2 = 0;
            end
end
for i3=0:1
    if i3==1
        f3 = state_list(i,4);
    else
        f3 = 0;
    end
for i4=0:1
    if i4==1
        f4 = state_list(i,5);
    else
        f4 = 0;
    end
for i5=0:1
    if i5==1
        f5 = state_list(i,6);
    else
        f5 = 0;
    end
for i6=0:1
    if i6==1
        f6 = state_list(i,7);
    else
        f6 = 0;
    end
f6 = 0;
end
actual = [A1 A2 A3 A4 A5 A6];
found = [f1 f2 f3 f4 f5 f6];
X = sort([f1 f2 f3 f4 f5 f6], 'descend');
F1 = X(1); F2 = X(2); F3 = X(3); F4 = X(4); F5 = X(5); F6 = X(6);
found_state = states(F1+1,F2+1,F3+1,F4+1,F5+1,F6+1);
p_f_a_temp = 1;
for j=1:n_trajectories
    if actual(j)==1
        sen_size = sen_s;
        sen_year_temp = 0;
        sen_prep_temp = sen_prep;
    elseif actual(j)==2
        sen_size = sen_m;
        sen_year_temp = 0;
        sen_prep_temp = sen_prep;
    elseif actual(j)==3
        sen_size = sen_l;
        sen_year_temp = 0;
        sen_prep_temp = 0;
    else

sen_size = sen_c;

sen_year_temp = 0;

sen_prep_temp = 0;

end

if actual(j) == 0
    sen = 1;
elseif actual(j) == found(j)
    sen = sen_size + sen_prep_temp + sen_year_temp;
    sen = exp(sen) / (1+exp(sen));
else
    sen = sen_size + sen_prep_temp + sen_year_temp;
    sen = exp(sen) / (1+exp(sen));
    sen = 1 - sen;
end

p_f_a_temp = p_f_a_temp * sen;
end

p_f_a(found_state, i, prep) = p_f_a(found_state, i, prep) + p_f_a_temp;
end
end

\[ p_{f\_a(:,i,\text{prep})} = \frac{p_{f\_a(:,i,\text{prep})}}{\text{sum}(p_{f\_a(:,i,\text{prep})})}; \]

end
20. COLONOSCOPY_6()

The colonoscopy function goes through each state and sees when a colonoscopy should be performed. It also removes any abnormalities found during a colonoscopy.

function [colonoscopy] = COLONOSCOPY_6()
global max_state state_list states
colonoscopy = zeros(max_state,max_state);
for i=1:max_state
    for i1=0:1
        if i1==0
            j1 = state_list(i,2) + 1;
            k1 = 1;
        else
            j1 = 1;
            k1 = state_list(i,2) + 1;
        end
        for i2=0:1
            if i2==0
                j2 = state_list(i,3) + 1;
                k2 = 1;
            else
                j2 = 1;
                k2 = state_list(i,3) + 1;
            end
\begin{verbatim}
end

for i3=0:1
    if i3==0
        j3 = state_list(i,4) + 1;
        k3 = 1;
    else
        j3 = 1;
        k3 = state_list(i,4) + 1;
    end

for i4=0:1
    if i4==0
        j4 = state_list(i,5) + 1;
        k4 = 1;
    else
        j4 = 1;
        k4 = state_list(i,5) + 1;
    end

for i5=0:1
    if i5==0
        j5 = state_list(i,6) + 1;
        k5 = 1;
    else
        j5 = 1;
\end{verbatim}
k5 = state_list(i,6) + 1;

end

for i6=0:1

    if i6==0

        j6 = state_list(i,7) + 1;
        k6 = 1;

    else

        j6 = 1;
        k6 = state_list(i,7) + 1;

    end

    K = sort([k6 k5 k4 k3 k2 k1],'descend');
    J = sort([j6 j5 j4 j3 j2 j1],'descend');

    K1 = K(1);  K2 = K(2);  K3 = K(3);  K4 = K(4);  K5 = K(5);  K6 = K(6);
    J1 = J(1);  J2 = J(2);  J3 = J(3);  J4 = J(4);  J5 = J(5);  J6 = J(6);

    remove = states(K1,K2,K3,K4,K5,K6);
    remain = states(J1,J2,J3,J4,J5,J6);

    colonoscopy(i,remove) = remain;

end

end

end

end
21. STATE_GROUP()

The state group function regroups the 924 states from STATE_DEFINITIONS into the 8 findings categories: (1) no adenomas, (2) 1-2 non-advance adenomas only, (3) 3+ non-advanced adenomas only, (4) 1-2 adenomas with some large/advanced adenomas, (5) 3+ adenomas with some large/advanced adenomas, (6) local CRC, (7) regional CRC, and, (8) distant CRC.

function [state_group] = STATE_GROUP()

  global max_state state_list n_states
  max_group = 8;

  % STATE_GROUP
  state_group = zeros(max_state, 2);
  for i=1:max_state
      state_group(i,1) = state_list(i,1);
      i1 = state_list(i,2);
      i2 = state_list(i,3);
      i3 = state_list(i,4);
      i4 = state_list(i,5);
      i5 = state_list(i,6);
      i6 = state_list(i,7);
      sm1 = 0;  sm2 = 0;  sm3 = 0;  sm4 = 0;  sm5 = 0;  sm6 = 0;
ad1 = 0;    ad2 = 0;    ad3 = 0;    ad4 = 0;    ad5 = 0;    ad6 = 0;
if i1==1 || i1==2
    sm1 = 1;
end
if i2==1 || i2==2
    sm2 = 1;
end
if i3==1 || i3==2
    sm3 = 1;
end
if i4==1 || i4==2
    sm4 = 1;
end
if i5==1 || i5==2
    sm5 = 1;
end
if i6==1 || i6==2
    sm6 = 1;
end
sm = sm1 + sm2 + sm3 + sm4 + sm5 + sm6;
if i1==1 || i1==2 || i1==3
    ad1 = 1;
end
if i2==1 || i2==2 || i2==3
    ad2 = 1;
end

if i3==1 || i3==2 || i3==3
    ad3 = 1;
end

if i4==1 || i4==2 || i4==3
    ad4 = 1;
end

if i5==1 || i5==2 || i5==3
    ad5 = 1;
end

if i6==1 || i6==2 || i6==3
    ad6 = 1;
end

ad = ad1 + ad2 + ad3 + ad4 + ad5 + ad6;

if i1==n_states || i2==n_states || i3==n_states || i4==n_states || i5==n_states ||
i6==n_states
    state_group(i,2) = 8;
elseif i1==(n_states-1) || i2==(n_states-1) || i3==(n_states-1) || i4==(n_states-1) ||
i5==(n_states-1) || i6==(n_states-1)
    state_group(i,2) = 7;
elseif i1==(n_states-2) || i2==(n_states-2) || i3==(n_states-2) || i4==(n_states-2) ||
    i5==(n_states-2) || i6==(n_states-2)
    state_group(i,2) = 6;
elseif i1==3 || i2==3 || i3==3 || i4==3 || i5==3 || i6==3
    if ad>=3
        state_group(i,2) = 5;
    else
        state_group(i,2) = 4;
    end
elseif i1==0 && i2==0 && i3==0 && i4==0 && i5==0 && i6==0
    state_group(i,2) = 1;
elseif sm<3
    state_group(i,2) = 2;
else
    state_group(i,2) = 3;
end
end
22. P_HEALTHY_DEATH ( )

In this function the probability of a healthy patient dying from non CRC reasons is calculated.

function [P_healthy_death] = P_HEALTHY_DEATH()

    global age_min age_max dt

    % mortality data
    data_ages = [2.5,7,12,17,22,25,32,37,42,47,52,57,62,67,72,77,82];
    data_deaths = [29.6,12.8,16.3,69.6,126.4,135.7,147.7,175.4,248.4,401,613.5,911.2,1269.2,1871.3,2831.9,4493.7,7358.2];
    data_deaths = data_deaths / 100000;

    % interpolate to get mortality probabilities
    interp_ages = age_min:dt:age_max;
    P_healthy_death = spline(data_ages,data_deaths,interp_ages);
23. P_CRC_DEATH()

In this function the probability of dying from the disease is given.

```matlab
function [P_local_death,P_regional_death,P_distant_death] = P_CRC_DEATH()

P_local_death(1) = 0.025120148;
P_local_death(2) = 0.025120148;
P_local_death(3) = 0.025120148;
P_local_death(4) = 0.025120148;
P_local_death(5) = 0.025120148;

P_regional_death(1) = 0.068866867;
P_regional_death(2) = 0.068866867;
P_regional_death(3) = 0.068866867;
P_regional_death(4) = 0.068866867;
P_regional_death(5) = 0.068866867;

P_distant_death(1) = 0.600507798;
P_distant_death(2) = 0.600507798;
P_distant_death(3) = 0.600507798;
P_distant_death(4) = 0.600507798;
P_distant_death(5) = 0.600507798;
```
24. P_HEALTHY_CRC()

In this function the probability of a healthy patient getting CRC is calculated. This function calculates the colon state as a person ages, adjusts the colon state when there is a colonoscopy and calculates the probability of developing cancer.

```matlab
function [P_healthy_CRC] = P_HEALTHY_CRC(N_ages,age_start)

P_healthy_CRC = zeros(N_ages,1);

[N_rows,N_cols,N_times] = size(K_markov);
N_states = N_rows;
P_state_current = zeros(N_rows,1);
P_state_current(1,1) = 1;
P_state = zeros(N_rows,N_times);
P_state(:,1) = P_state_current;
P_group_current = zeros(6,1);
P_group_current(1,1) = 1;
P_group = zeros(6,N_times);
P_group(:,1) = P_group_current;
for T=2:N_times
    % calculate colon state as a person ages
    P_state_next = zeros(N_rows,1);
P_transition = K_markov(:,:,T);
```

for Finish_state=1:N_rows
    for Start_state=1:N_cols
        P_state_next(Finish_state) = P_state_next(Finish_state) +
        P_transition(Finish_state,Start_state) * P_state_current(Start_state);
    end
end

P_state(:,T) = P_state_next;
P_state_current = P_state_next;
P_group_current = zeros(6,1);
for State=1:N_states
    Group = state_group(State,2);
    P_group_current(Group) = P_group_current(Group) +
    P_state_current(State);
end

P_group(:,T) = P_group_current;

% adjust colon state when there is a colonoscopy
if T==age_start
    P_state_observed = zeros(N_rows,1);
P_group_observed=zeros(6,1);
    for Observed_state=1:N_rows
        for Actual_state=1:N_cols
            P_state_observed(Observed_state) = P_state_observed(Observed_state) +
p_f_a(Observed_state,Actual_state,2,10) * P_state(Actual_state,T);
for State=1:N_states
    Group = state_group(State,2);
    P_group_observed(Group) = P_group_observed(Group) + P_state_observed(State);
end

Post_colonoscopy=zeros(N_rows,1);
for Actual_state=1:N_rows
    for Observed_state=1:N_cols
        Resulting_state = colonoscopy(Actual_state,Observed_state);
        if Resulting_state~=0
            P_actual_state = P_state(Actual_state,T);
            P_observed_state = p_f_a(Observed_state,Actual_state,2,10);
            Post_colonoscopy(Resulting_state) = Post_colonoscopy(Resulting_state) + P_actual_state * P_observed_state;
        end
    end
end
P_state_current = Post_colonoscopy;
end

% probability of developing cancer
if T==(age_start+1)
P_healthy_CRC(T) = P_group(6,T);

elseif T>(age_start+1)

    P_healthy_CRC(T) = P_group(6,T) - P_group(6,T-1);

end

definitions
function [state_transitions X_diff] = STATE_TRANSITIONS()

global n_trajectories
global state_list
global max_state

state_transitions = zeros(max_state, max_state);
X_diff = zeros(max_state, max_state);
for state_start = 1:max_state
    for state_finish = 1:max_state
        if state_start ~= state_finish
            n_diff = 0;
            n_diff_1 = 0;
            n_diff_24 = 0;
            n_diff_30 = 0;
            n_diff_20 = 0;
            n_diff_10 = 0;
            for i=1:n_trajectories
                I = i + 1;
                if state_list(state_start,I)~=state_list(state_finish,I)
                    n_diff = n_diff + 1;
                end
            end
        end
    end
end
if state_list(state_start,I)+1==state_list(state_finish,I)
    n_diff_1 = n_diff_1 + 1;
    I_diff = I;
    x_diff = 1;
    y_diff = 1;
    if state_list(state_start,I)>0
        x_diff = x_diff - 1;
        for j=1:n_trajectories
            J = j + 1;
            if state_list(state_start,J)==state_list(state_start,I)
                x_diff = x_diff + 1;
            end
        end
    end
    if state_list(state_finish,I)>0
        y_diff = y_diff - 1;
        for j=1:n_trajectories
            J = j + 1;
            if state_list(state_finish,J)==state_list(state_start,I)
                y_diff = y_diff + 1;
            end
        end
    end
end
elseif state_list(state_start,I)==2 && state_list(state_finish,I)==4

    n_diff_24 = n_diff_24 + 1;
    I_diff = I;
    x_diff = 1;
    y_diff = 1;
    if state_list(state_start,I)>0
        x_diff = x_diff - 1;
        for j=1:n_trajectories
            J = j + 1;
            if state_list(state_start,J)==state_list(state_start,I)
                x_diff = x_diff + 1;
            end
        end
    end
end

elseif state_list(state_start,I)==3 && state_list(state_finish,n_trajectories+1)==0

    go = 0;
    if I==2 &&
        sum(state_list(state_start,I+1:n_trajectories+1)==state_list(state_finish,I:n_trajectories+1))/length(I:n_trajectories+1)==1
        go = 1;
        elseif I==n_trajectories+1 && sum(state_list(state_start,2:I-1)==state_list(state_finish,2:I-1))/length(2:I-1)==1
            go = 1;
        end
end
go = 1;

elseif sum(state_list(state_start,2:I-1)==state_list(state_finish,2:I-1))/length(2:I-1)==1 &&
sum(state_list(state_start,I+1:n_trajectories+1)==state_list(state_finish,I:n_trajectories))/length(I:n_trajectories)==1

go = 1;

end

if go==1

n_diff_30 = n_diff_30 + 1;

if n_diff_30==1

I_diff = 1;

x_diff = 1;

y_diff = 1;

if state_list(state_start,I)>0

x_diff = x_diff - 1;

for j=1:n_trajectories

J = j + 1;

if state_list(state_start,J)==state_list(state_start,I)

x_diff = x_diff + 1;

end

end

end

end
end

elseif state_list(state_start,I)==2 &&
state_list(state_finish,n_trajectories+1)==0
    go = 0;
    if I==2 &&
    sum(state_list(state_start,I+1:n_trajectories+1)==state_list(state_finish,I:n_trajectories))/length(I:n_trajectories)==1
        go = 1;
    elseif I==n_trajectories+1 && sum(state_list(state_start,2:I-1)==state_list(state_finish,2:I-1))/length(2:I-1)==1
        go = 1;
    elseif sum(state_list(state_start,2:I-1)==state_list(state_finish,2:I-1))/length(2:I-1)==1 &&
    sum(state_list(state_start,I+1:n_trajectories+1)==state_list(state_finish,I:n_trajectories))/length(I:n_trajectories)==1
        go = 1;
    end

if go==1
    n_diff_20 = n_diff_20 + 1;
    if n_diff_20==1
        l_diff = I;
        x_diff = 1;
        y_diff = 1;
    end
if state_list(state_start,I)>0
    x_diff = x_diff - 1;
    for j=1:n_trajectories
        J = j + 1;
        if state_list(state_start,J)==state_list(state_start,I)
            x_diff = x_diff + 1;
        end
    end
end

elseif state_list(state_start,I)==1 && state_list(state_finish,n_trajectories+1)==0
    go = 0;
    if I==2 && sum(state_list(state_start,I+1:n_trajectories+1)==state_list(state_finish,I:n_trajectories))/length(I:n_trajectories)==1
        go = 1;
    elseif I==n_trajectories+1 && sum(state_list(state_start,2:I-1)==state_list(state_finish,2:I-1))/length(2:I-1)==1
        go = 1;
    elseif sum(state_list(state_start,2:I-1)==state_list(state_finish,2:I-1))/length(2:I-1)==1 &&
sum(state_list(state_start,I+1:n_trajectories+1)==state_list(state_finish,I:n_trajectories))/length(I:n_trajectories)==1

    go = 1;

end

if go==1

    n_diff_10 = n_diff_10 + 1;

    if n_diff_10==1

        I_diff = I;

        x_diff = 1;

        y_diff = 1;

        if state_list(state_start,I)>0

            x_diff = x_diff - 1;

            for j=1:n_trajectories

                J = j + 1;

                if state_list(state_start,J)==state_list(state_start,I)

                    x_diff = x_diff + 1;

                end

            end

        end

    end

end

end

end
end

if n_diff_30>=1

    state_transitions(state_finish,state_start) = 30;
    X_diff(state_finish,state_start) = x_diff;

elseif n_diff_20>=1

    state_transitions(state_finish,state_start) = 20;
    X_diff(state_finish,state_start) = x_diff;

elseif n_diff_10>=1

    state_transitions(state_finish,state_start) = 10;
    X_diff(state_finish,state_start) = x_diff;

elseif n_diff==1 && n_diff_24==1 && state_list(state_start,I_diff)==2

    && state_list(state_finish,I_diff)==4

    state_transitions(state_finish,state_start) = 24;
    X_diff(state_finish,state_start) = x_diff;

elseif n_diff==1 && n_diff_1==1

    if state_list(state_start,I_diff)==0

        state_transitions(state_finish,state_start) = 1;
        X_diff(state_finish,state_start) = x_diff;

    elseif state_list(state_start,I_diff)==1

        state_transitions(state_finish,state_start) = 12;
        X_diff(state_finish,state_start) = x_diff;

        state_transitions(state_start,state_finish) = 21;
        X_diff(state_start,state_finish) = y_diff;
elseif state_list(state_start,I_diff)==2
    state_transitions(state_finish,state_start) = 23;
    X_diff(state_finish,state_start) = x_diff;
    state_transitions(state_start,state_finish) = 32;
    X_diff(state_start,state_finish) = y_diff;
elseif state_list(state_start,I_diff)==3
    state_transitions(state_finish,state_start) = 34;
    X_diff(state_finish,state_start) = x_diff;
end
end
end
end
end
26. STATE_DEFINITIONS ( )

This function gives all the combination of states.

function [states, state_list, max_state] = STATE_DEFINITIONS()

global n_states

states =
zeros(n_states+1,n_states+1,n_states+1,n_states+1,n_states+1,n_states+1);
state = 0;
for i1=0:n_states
    for i2=0:n_states
        for i3=0:n_states
            for i4=0:n_states
                for i5=0:n_states
                    for i6=0:n_states
                        if (i1>=i2 && i2>=i3 && i3>=i4 && i4>=i5 && i5>=i6)
                            state = state + 1;
                            states(i1+1,i2+1,i3+1,i4+1,i5+1,i6+1) = state;
                            state_list(state,:) = [state i1 i2 i3 i4 i5 i6];
                        end
                    end
                end
            end
        end
    end
end
max_state = state
27. TRANSITION_RATES

In this function the rate at which adenoma forms and progresses is calculated. Symptom development rates are also developed in this function.

function [gammas] = TRANSITION_RATES(ks)

global n_ts

global n_transitions

global dtau t_max

gammas = zeros(n_ts,n_transitions+1);

% adenoma formation
taus = 0:dtau:t_max;
gammas(:,1) = max(0,spline(0:10:t_max,[0 ks(1,:)],taus));

% adenoma progression
for i = 2:n_transitions
    gammas(:,i) = max(0,spline(0:10:t_max,[0 ks(i,:)],taus));
end

% symptom development

% gammas(:,n_transitions+1) = max(0,spline(0:10:t_max,[0 ks(n_transitions+1,:)],[tau]));
\[ \text{gammas}(;n_{\text{transitions}}+2) = \max(0, \text{spline}(0:10:t_{\text{max}},[0 \text{ks}(n_{\text{transitions}}+2,:)], \text{taus})); \]

\[ \text{gammas}(;n_{\text{transitions}}+3) = \max(0, \text{spline}(0:10:t_{\text{max}},[0 \text{ks}(n_{\text{transitions}}+3,:)], \text{taus})); \]
28. TREATMENT_DEFINITIONS

This function gives the combination of positive and negative FITs. With FITs order matters. There are a total of 127 FIT treatments.

function [treatments, treatment_list, max_treatment] = TREATMENT_DEFINITIONS_7()

  treatments = zeros(3,3,3,3,3,3,3,3);
  treatment = 0;
  for i7=0:2
    for i6=0:2
      for i5=0:2
        for i4=0:2
          for i3=0:2
            for i2=0:2
              for i1=0:2
                if i1==2
                  i1_temp = -1;
                else
                  i1_temp = i1;
                end
                if i2==2
                  i2_temp = -1;
                else
                  i2_temp = i2;
                end
                if i3==2
                  i3_temp = -1;
                else
                  i3_temp = i3;
                end
                if i4==2
                  i4_temp = -1;
                else
                  i4_temp = i4;
                end
                if i5==2
                  i5_temp = -1;
                else
                  i5_temp = i5;
                end
                if i6==2
                  i6_temp = -1;
                else
                  i6_temp = i6;
                end
              end
            end
          end
        end
      end
    end
  end

i6_temp = i6;
end
if i7==2
    i7_temp = -1;
else
    i7_temp = i7;
end

if (abs(i1_temp)>=abs(i2_temp) && abs(i2_temp)>=abs(i3_temp) &
    && abs(i3_temp)>=abs(i4_temp) && abs(i4_temp)>=abs(i5_temp) &&
    && abs(i5_temp)>=abs(i6_temp) && abs(i6_temp)>=abs(i7_temp))

    treatment = treatment + 1;
treatments(i1_temp+2,i2_temp+2,i3_temp+2,i4_temp+2,i5_temp+2,i6_temp+2,i7_temp+2) = treatment;
    treatment_list(treatment,:) = [treatment i1_temp i2_temp i3_temp i4_temp i5_temp i6_temp i7_temp];
end
end
end
end
end
end
end
max_treatment = treatment;
29. P_Patient_COOLON_FIT_combo

This function calculates the probabilities from age 0 to death or age 100. This function has a series of blocks computing different probabilities. It calculates healthy mortality for the year, cancer related mortality, colon state as a person ages, and the probability of colonoscopy based on symptom development. In this function is where patients get a screening colonoscopy then followed by a FIT follow-up. It also calculates the yearly probabilities at each block.

P_Patient_COOLON_FIT_combo(T1,positive_step,negative_step,stop_age)

function [P_death, P_Asymptomatic, P_local, P_regional, P_distant, P_colonoscopy, P_FIT] =

P_Patient_COOLON_FIT_combo(T1,positive_step,negative_step,stop_age)

global total_ages

global K_markov

global P_local_death P_regional_death P_distant_death

global P_healthy_death

global state_list max_state

global p_f_a colonoscopy state_group

max_treatment = 255;
start_age = T1;

[FIT_AGE_matrix] = FIT_AGE (start_age, positive_step, negative_step);

[fit_matrix] = FIT_MATRIX ();

FIT_AGE_matrix(:,2) = FIT_AGE_matrix(:,2) + positive_step;

P_colonoscopy = zeros(total_ages,1);

P_FIT = zeros(total_ages,1);

P_Asymptomatic = zeros(total_ages,1);

P_Asymptomatic(1) = 1;

P_local = zeros(total_ages,5);

P_regional = zeros(total_ages,5);

P_distant = zeros(total_ages,5);

P_death = zeros(total_ages,1);

[N_rows,N_cols,N_times] = size(K_markov);

P_state = zeros(N_rows,max_treatment);

P_state(1,1) = 1;

P_state_local = zeros(N_rows,5,max_treatment);

P_state_regional = zeros(N_rows,5,max_treatment);

P_state_distant = zeros(N_rows,5,max_treatment);
for T=2:(T1+1)
    age = T - 1;
    current_treatment = 1;
    age_index = T ;
    P_death(age_index) = P_death(age_index-1);

    % healthy mortality during the year
    for state=1:N_rows
        P_death(age_index) = P_death(age_index) + P_healthy_death(age_index) * 
        P_state(state,current_treatment);
        P_state(state,current_treatment) = P_state(state,current_treatment) - 
        P_healthy_death(age_index) * P_state(state,current_treatment);
    end

    % cancer related mortality during the year
    for i=1:5
        P_death(age_index) = P_death(age_index) + max(P_local_death(i),
        P_healthy_death(age_index)) * sum(P_state_local(:,i,current_treatment));
    end
P_state_local(:,i,current_treatment) = P_state_local(:,i,current_treatment) -
max(P_local_death(i), P_healthy_death(age_index)) * 
P_state_local(:,i,current_treatment);

P_death(age_index) = P_death(age_index) + max(P_regional_death(i),
P_healthy_death(age_index)) * sum(P_state_regional(:,i,current_treatment));

P_state_regional(:,i,current_treatment) =
P_state_regional(:,i,current_treatment) - max(P_regional_death(i),
P_healthy_death(age_index)) * P_state_regional(:,i,current_treatment);

P_death(age_index) = P_death(age_index) + max(P_distant_death(i),
P_healthy_death(age_index)) * sum(P_state_distant(:,i,current_treatment));

P_state_distant(:,i,current_treatment) = P_state_distant(:,i,current_treatment) - max(P_distant_death(i), P_healthy_death(age_index)) * 
P_state_distant(:,i,current_treatment);

end

% calculate colon state as a person ages
P_state_next = zeros(N_rows,1);
P_state_next_local = zeros(N_rows,5);
P_state_next_regional = zeros(N_rows,5);
P_state_next_distant = zeros(N_rows,5);
P_transition = K_markov(:,:,age_index);
for Finish_state=1:N_rows
    for Start_state=1:N_cols
\[ P_{\text{state\_next}}(\text{Finish\_state}) = P_{\text{state\_next}}(\text{Finish\_state}) + P_{\text{transition}}(\text{Finish\_state},\text{Start\_state}) \times P_{\text{state}}(\text{Start\_state}, \text{current\_treatment}); \]

\[
\text{for } i=1:5
\]

\[ P_{\text{state\_next\_local}}(\text{Finish\_state},i) = P_{\text{state\_next\_local}}(\text{Finish\_state},i) + P_{\text{transition}}(\text{Finish\_state},\text{Start\_state}) \times P_{\text{state\_local}}(\text{Start\_state},i,\text{current\_treatment}); \]

\[ P_{\text{state\_next\_regional}}(\text{Finish\_state},i) = P_{\text{state\_next\_regional}}(\text{Finish\_state},i) + P_{\text{transition}}(\text{Finish\_state},\text{Start\_state}) \times P_{\text{state\_regional}}(\text{Start\_state},i,\text{current\_treatment}); \]

\[ P_{\text{state\_next\_distant}}(\text{Finish\_state},i) = P_{\text{state\_next\_distant}}(\text{Finish\_state},i) + P_{\text{transition}}(\text{Finish\_state},\text{Start\_state}) \times P_{\text{state\_distant}}(\text{Start\_state},i,\text{current\_treatment}); \]

\[
\text{end}
\]

\[
\text{end}
\]

\[
\text{end}
\]

\[
\text{for } i=1:4
\]

\[ P_{\text{state\_local}}(:,i+1,\text{current\_treatment}) = P_{\text{state\_next\_local}}(:,i); \]

\[ P_{\text{state\_regional}}(:,i+1,\text{current\_treatment}) = P_{\text{state\_next\_regional}}(:,i); \]

\[ P_{\text{state\_distant}}(:,i+1,\text{current\_treatment}) = P_{\text{state\_next\_distant}}(:,i); \]

\[
\text{end}
\]

\[ P_{\text{state}}(:,\text{current\_treatment}) = P_{\text{state\_next}} + P_{\text{state\_next\_local}}(:,5) + P_{\text{state\_next\_regional}}(:,5) + P_{\text{state\_next\_distant}}(:,5); \]

% colonoscopy for symptoms
\[ P_{\text{state\_post}} = P_{\text{state}}(:,\text{current\_treatment}); \]
\[ P_{\text{state\_local}}(:,1,\text{current\_treatment}) = \text{zeros}(N_{\text{rows}},1,1); \]
\[ P_{\text{state\_regional}}(:,1,\text{current\_treatment}) = \text{zeros}(N_{\text{rows}},1,1); \]
\[ P_{\text{state\_distant}}(:,1,\text{current\_treatment}) = \text{zeros}(N_{\text{rows}},1,1); \]

\begin{verbatim}
for Actual\_state=1:N\_rows
    P\_asym\_sym = 0;
    for i=6:-1:1
        if state\_list(Actual\_state,i+1)==4
            P\_asym\_sym = 0.17;
        elseif state\_list(Actual\_state,i+1)==5
            P\_asym\_sym = 0.22;
        elseif state\_list(Actual\_state,i+1)==6
            P\_asym\_sym = 0.50;
        end
    end

    P\_colonoscopy(age\_index) = P\_colonoscopy(age\_index) + P\_asym\_sym * P\_state(Actual\_state,\text{current\_treatment});

    for Observed\_state=1:N\_rows
        Group = state\_group(Observed\_state,2);
        Resulting\_state = colonoscopy(Actual\_state,Observed\_state);
        if Resulting\_state\_==0
            if Group\_==6

end
\end{verbatim}
\begin{align*}
\text{P}_\text{state\_local}(\text{Resulting\_state},1,\text{current\_treatment}) &= \\
&\text{P}_\text{state\_local}(\text{Resulting\_state},1,\text{current\_treatment}) + \text{P\_asym\_sym} \times \\
p\_f\_a(\text{Observed\_state},\text{Actual\_state},2) \times \text{P\_state}(\text{Actual\_state},\text{current\_treatment}) ; \\
\text{P\_state\_post}(\text{Actual\_state}) &= \text{P\_state\_post}(\text{Actual\_state}) - \\
\text{P\_asym\_sym} \times p\_f\_a(\text{Observed\_state},\text{Actual\_state},2) \times \\
\text{P\_state}(\text{Actual\_state},\text{current\_treatment}) ; \\
\text{elseif} \text{ Group}==7 \\
\text{P\_state\_regional}(\text{Resulting\_state},1,\text{current\_treatment}) &= \\
\text{P\_state\_regional}(\text{Resulting\_state},1,\text{current\_treatment}) + \text{P\_asym\_sym} \times \\
p\_f\_a(\text{Observed\_state},\text{Actual\_state},2) \times \text{P\_state}(\text{Actual\_state},\text{current\_treatment}) ; \\
\text{P\_state\_post}(\text{Actual\_state}) &= \text{P\_state\_post}(\text{Actual\_state}) - \\
\text{P\_asym\_sym} \times p\_f\_a(\text{Observed\_state},\text{Actual\_state},2) \times \\
\text{P\_state}(\text{Actual\_state},\text{current\_treatment}) ; \\
\text{elseif} \text{ Group}==8 \\
\text{P\_state\_distant}(\text{Resulting\_state},1,\text{current\_treatment}) &= \\
\text{P\_state\_distant}(\text{Resulting\_state},1,\text{current\_treatment}) + \text{P\_asym\_sym} \times \\
p\_f\_a(\text{Observed\_state},\text{Actual\_state},2) \times \text{P\_state}(\text{Actual\_state},\text{current\_treatment}) ; \\
\text{P\_state\_post}(\text{Actual\_state}) &= \text{P\_state\_post}(\text{Actual\_state}) - \\
\text{P\_asym\_sym} \times p\_f\_a(\text{Observed\_state},\text{Actual\_state},2) \times \\
\text{P\_state}(\text{Actual\_state},\text{current\_treatment}) ; \\
\text{else} 
\end{align*}
\[ P_{\text{state\_post}}(\text{Resulting\_state}) = P_{\text{state\_post}}(\text{Resulting\_state}) + \]
\[ P_{\text{asym\_sym}} \times p_{f \_a}(\text{Observed\_state}, \text{Actual\_state}, 2) \times \]
\[ P_{\text{state}}(\text{Actual\_state, current\_treatment}); \]
\[ P_{\text{state\_post}}(\text{Actual\_state}) = P_{\text{state\_post}}(\text{Actual\_state}) - \]
\[ P_{\text{asym\_sym}} \times p_{f \_a}(\text{Observed\_state}, \text{Actual\_state}, 2) \times \]
\[ P_{\text{state}}(\text{Actual\_state, current\_treatment}); \]
\] 
\end{end}
\end{end}
\end{end}
\end{end}
P_{\text{state}}(:,\text{current\_treatment}) = P_{\text{state\_post}};

% yearly probabilities
\[
P_{\text{Asymptomatic}}(\text{age\_index}) = \text{sum(sum}(P_{\text{state}}));\]
\[
\text{for} \ i=1:5 \]
\[
P_{\text{local}}(\text{age\_index},i) = \text{sum(sum}(P_{\text{state\_local}}(:,i,:)));\]
\[
P_{\text{regional}}(\text{age\_index},i) = \text{sum(sum}(P_{\text{state\_regional}}(:,i,:)));\]
\[
P_{\text{distant}}(\text{age\_index},i) = \text{sum(sum}(P_{\text{state\_distant}}(:,i,:)));\]
\[
\text{end} \]
\[
\text{if} \ T\sim(T+1) \]
\[
\text{age} \]
100 * (P_death(age_index) + P_Asymptomatic(age_index) + 
sum(P_local(age_index,:)) + sum(P_regional(age_index,:)) + 
sum(P_distant(age_index,:)))

100 * [P_death(age_index); P_Asymptomatic(age_index); 
sum(P_local(age_index,:)); sum(P_regional(age_index,:)); sum(P_distant(age_index,:)); 
P_colonoscopy(age_index); P_FIT(age_index)]

end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% 2nd block - screening colonoscopy
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

P_colonoscopy(age_index) = sum(P_state(:,current_treatment));
P_state_post = P_state(:,current_treatment);
for Observed_state=1:N_rows
    Group = state_group(Observed_state,2);
    for Actual_state=1:N_cols
        Resulting_state = colonoscopy(Actual_state,Observed_state);
        if Resulting_state~=0
            if Group==6
                P_state_local(Resulting_state,1,current_treatment) = 
P_state_local(Resulting_state,1,current_treatment) + P_asym_sym *
p_f_a(Observed_state,Actual_state,2) * P_state(Actual_state,current_treatment);
\[ P_{\text{state\_post}}(\text{Actual\_state}) = P_{\text{state\_post}}(\text{Actual\_state}) - P_{\text{sym\_sym}} \times p_{\text{f\_a}}(\text{Observed\_state}, \text{Actual\_state}, 2) \times P_{\text{state}}(\text{Actual\_state}, \text{current\_treatment}); \]

\[ \text{elseif Group} == 7 \]

\[ P_{\text{state\_regional}}(\text{Resulting\_state}, 1, \text{current\_treatment}) = P_{\text{state\_regional}}(\text{Resulting\_state}, 1, \text{current\_treatment}) + P_{\text{sym\_sym}} \times p_{\text{f\_a}}(\text{Observed\_state}, \text{Actual\_state}, 2) \times P_{\text{state}}(\text{Actual\_state}, \text{current\_treatment}); \]

\[ P_{\text{state\_post}}(\text{Actual\_state}) = P_{\text{state\_post}}(\text{Actual\_state}) - P_{\text{sym\_sym}} \times p_{\text{f\_a}}(\text{Observed\_state}, \text{Actual\_state}, 2) \times P_{\text{state}}(\text{Actual\_state}, \text{current\_treatment}); \]

\[ \text{elseif Group} == 8 \]

\[ P_{\text{state\_distant}}(\text{Resulting\_state}, 1, \text{current\_treatment}) = P_{\text{state\_distant}}(\text{Resulting\_state}, 1, \text{current\_treatment}) + P_{\text{sym\_sym}} \times p_{\text{f\_a}}(\text{Observed\_state}, \text{Actual\_state}, 2) \times P_{\text{state}}(\text{Actual\_state}, \text{current\_treatment}); \]

\[ P_{\text{state\_post}}(\text{Actual\_state}) = P_{\text{state\_post}}(\text{Actual\_state}) - P_{\text{sym\_sym}} \times p_{\text{f\_a}}(\text{Observed\_state}, \text{Actual\_state}, 2) \times P_{\text{state}}(\text{Actual\_state}, \text{current\_treatment}); \]

\[ \text{else} \]

\[ P_{\text{state\_post}}(\text{Resulting\_state}) = P_{\text{state\_post}}(\text{Resulting\_state}) + P_{\text{sym\_sym}} \times p_{\text{f\_a}}(\text{Observed\_state}, \text{Actual\_state}, 2) \times P_{\text{state}}(\text{Actual\_state}, \text{current\_treatment}); \]

\[ P_{\text{state\_post}}(\text{Actual\_state}) = P_{\text{state\_post}}(\text{Actual\_state}) - P_{\text{sym\_sym}} \times p_{\text{f\_a}}(\text{Observed\_state}, \text{Actual\_state}, 2) \times P_{\text{state}}(\text{Actual\_state}, \text{current\_treatment}); \]

\[ \text{end} \]

\[ \text{end} \]

\[ \text{end} \]
P_state(:,current_treatment) = P_state_post;

% yearly probabilities
P_Asymptomatic(age_index) = sum(sum(P_state));
for i=1:5
    P_local(age_index,i) = sum(sum(P_state_local(:,i,:)));  
    P_regional(age_index,i) = sum(sum(P_state_regional(:,i,:)));  
    P_distant(age_index,i) = sum(sum(P_state_distant(:,i,:)));  
end

age

100 * (P_death(age_index) + P_Asymptomatic(age_index) +
sum(P_local(age_index,:)) + sum(P_regional(age_index,:)) +
sum(P_distant(age_index,:))))

100 * [P_death(age_index); P_Asymptomatic(age_index);
sum(P_local(age_index,:)); sum(P_regional(age_index,:)); sum(P_distant(age_index,:));
P_colonoscopy(age_index); P_FIT(age_index)]

% 3rd block - FIT after colonoscopy %%%

% FIT after colonoscopy
for T=(T1+2):N_times
    age = T - 1;
end
age\textunderscore index = T;

P\textunderscore death(age\textunderscore index) = P\textunderscore death(age\textunderscore index-1);

for current\textunderscore treatment=max\textunderscore treatment:-1:1

if sum(P\textunderscore state(:,current\textunderscore treatment)) > 0

% healthy mortality during the year

for state=1:N\_rows

P\textunderscore death(age\textunderscore index) = P\textunderscore death(age\textunderscore index) +
P\textunderscore healthy\textunderscore death(age\textunderscore index) * P\textunderscore state(state,current\textunderscore treatment);

P\textunderscore state(state,current\textunderscore treatment) = P\textunderscore state(state,current\textunderscore treatment) -
P\textunderscore healthy\textunderscore death(age\textunderscore index) * P\textunderscore state(state,current\textunderscore treatment);

end

% cancer related mortality during the year

for i=1:5

P\textunderscore death(age\textunderscore index) = P\textunderscore death(age\textunderscore index) + max(P\textunderscore local\textunderscore death(i),
P\textunderscore healthy\textunderscore death(age\textunderscore index)) * sum(P\textunderscore state\textunderscore local(:,i,current\textunderscore treatment));

P\textunderscore state\textunderscore local(:,i,current\textunderscore treatment) =
P\textunderscore state\textunderscore local(:,i,current\textunderscore treatment) - max(P\textunderscore local\textunderscore death(i), P\textunderscore healthy\textunderscore death(age\textunderscore index))
* P\textunderscore state\textunderscore local(:,i,current\textunderscore treatment);
\[
P_{\text{death}}(\text{age\_index}) = P_{\text{death}}(\text{age\_index}) + \max(P_{\text{regional\_death}}(i), \nonumber \\
P_{\text{healthy\_death}}(\text{age\_index})) \ast \sum(P_{\text{state\_regional}}(:,i,\text{current\_treatment})); \nonumber \\
P_{\text{state\_regional}}(:,i,\text{current\_treatment}) = \\
P_{\text{state\_regional}}(:,i,\text{current\_treatment}) - \max(P_{\text{regional\_death}}(i), \nonumber \\
P_{\text{healthy\_death}}(\text{age\_index})) \ast P_{\text{state\_regional}}(:,i,\text{current\_treatment}); \nonumber \\
P_{\text{death}}(\text{age\_index}) = P_{\text{death}}(\text{age\_index}) + \max(P_{\text{distant\_death}}(i), \nonumber \\
P_{\text{healthy\_death}}(\text{age\_index})) \ast \sum(P_{\text{state\_distant}}(:,i,\text{current\_treatment})); \nonumber \\
P_{\text{state\_distant}}(:,i,\text{current\_treatment}) = \\
P_{\text{state\_distant}}(:,i,\text{current\_treatment}) - \max(P_{\text{distant\_death}}(i), \nonumber \\
P_{\text{healthy\_death}}(\text{age\_index})) \ast P_{\text{state\_distant}}(:,i,\text{current\_treatment}); \nonumber \\
\end{align*}

End

\% calculate colon state as a person ages

\[
P_{\text{state\_next}} = \text{zeros}(\text{N\_rows},1); \nonumber \\
P_{\text{state\_next\_local}} = \text{zeros}(\text{N\_rows},5); \nonumber \\
P_{\text{state\_next\_regional}} = \text{zeros}(\text{N\_rows},5); \nonumber \\
P_{\text{state\_next\_distant}} = \text{zeros}(\text{N\_rows},5); \nonumber \\
P_{\text{transition}} = K_{\text{markov}}(:,:,\text{age\_index}); \nonumber \\
\end{align*}

for Finish$\_state$=1:N$\_rows$

for Start$\_state$=1:N$\_cols$

\[
P_{\text{state\_next}}(\text{Finish\_state}) = P_{\text{state\_next}}(\text{Finish\_state}) + \nonumber \\
P_{\text{transition}}(\text{Finish\_state},\text{Start\_state}) \ast P_{\text{state}}(\text{Start\_state}, \text{current\_treatment}); \nonumber \\
\end{align*}

for i=1:5
\begin{align*}
\text{P\_state\_next\_local}(\text{Finish\_state},i) &= \text{P\_state\_next\_local}(\text{Finish\_state},i) + \text{P\_transition}(\text{Finish\_state},\text{Start\_state}) \times \\
& \quad \text{P\_state\_local}(\text{Start\_state},i,\text{current\_treatment}); \\
\text{P\_state\_next\_regional}(\text{Finish\_state},i) &= \text{P\_state\_next\_regional}(\text{Finish\_state},i) + \text{P\_transition}(\text{Finish\_state},\text{Start\_state}) \times \\
& \quad \text{P\_state\_regional}(\text{Start\_state},i,\text{current\_treatment}); \\
\text{P\_state\_next\_distant}(\text{Finish\_state},i) &= \text{P\_state\_next\_distant}(\text{Finish\_state},i) + \text{P\_transition}(\text{Finish\_state},\text{Start\_state}) \times \\
& \quad \text{P\_state\_distant}(\text{Start\_state},i,\text{current\_treatment}); \\
\text{end} \quad \text{end} \quad \text{end}
\end{align*}

for i=1:4

\begin{align*}
\text{P\_state\_local}(:,i+1,\text{current\_treatment}) &= \text{P\_state\_next\_local}(:,i); \\
\text{P\_state\_regional}(:,i+1,\text{current\_treatment}) &= \text{P\_state\_next\_regional}(:,i); \\
\text{P\_state\_distant}(:,i+1,\text{current\_treatment}) &= \text{P\_state\_next\_distant}(:,i); \\
\text{end}
\end{align*}

\begin{align*}
\text{P\_state}(:,\text{current\_treatment}) &= \text{P\_state\_next} + \text{P\_state\_next\_local}(;,5) + \\
& \quad \text{P\_state\_next\_regional}(;,5) + \text{P\_state\_next\_distant}(;,5); \\
\text{% colonoscopy for symptoms}
\text{P\_state\_post} &= \text{P\_state}(:,\text{current\_treatment}); \\
\text{P\_state\_local}(;,1,\text{current\_treatment}) &= \text{zeros}(\text{N\_rows},1,1);
\end{align*}
\[
P_{\text{state\_regional}}(:,1,\text{current\_treatment}) = \text{zeros}(N_{\text{rows}},1,1);
\]

\[
P_{\text{state\_distant}}(:,1,\text{current\_treatment}) = \text{zeros}(N_{\text{rows}},1,1);
\]

for Actual\_state=1:N\_rows

\[
P_{\text{asym\_sym}} = 0;
\]

for i=6:-1:1

\[
\text{if state\_list(Actual\_state,i+1)==4}
\]

\[
P_{\text{asym\_sym}} = 0.17;
\]

\[
\text{elseif state\_list(Actual\_state,i+1)==5}
\]

\[
P_{\text{asym\_sym}} = 0.22;
\]

\[
\text{elseif state\_list(Actual\_state,i+1)==6}
\]

\[
P_{\text{asym\_sym}} = 0.50;
\]

end

end

\[
P_{\text{colonoscopy(age\_index)}} = P_{\text{colonoscopy(age\_index)}} + P_{\text{asym\_sym}} * P_{\text{state(Actual\_state, current\_treatment)}};
\]

for Observed\_state=1:N\_rows

\[
\text{Group = state\_group(Observed\_state,2)};
\]

\[
\text{Resulting\_state = colonoscopy(Actual\_state, Observed\_state)};
\]

\[
\text{if Resulting\_state==0}
\]

\[
\text{if Group==6}
\]

\[
P_{\text{state\_local(Resulting\_state,1, current\_treatment)}} = P_{\text{state\_local(Resulting\_state,1, current\_treatment)}} + P_{\text{asym\_sym}} * p_{f\_a(Observed\_state,Actual\_state,2)} * P_{\text{state(Actual\_state, current\_treatment)}};
\]
P_state_post(Actual_state) = P_state_post(Actual_state) -
P_asym_sym * p_f_a(Observed_state,Actual_state,2) *
P_state(Actual_state,current_treatment);

elseif Group==7

P_state_regional(Resulting_state,1,current_treatment) =
P_state_regional(Resulting_state,1,current_treatment) + P_asym_sym *
p_f_a(Observed_state,Actual_state,2) * P_state(Actual_state,current_treatment);

P_state_post(Actual_state) = P_state_post(Actual_state) -
P_asym_sym * p_f_a(Observed_state,Actual_state,2) *
P_state(Actual_state,current_treatment);

elseif Group==8

P_state_distant(Resulting_state,1,current_treatment) =
P_state_distant(Resulting_state,1,current_treatment) + P_asym_sym *
p_f_a(Observed_state,Actual_state,2) * P_state(Actual_state,current_treatment);

P_state_post(Actual_state) = P_state_post(Actual_state) -
P_asym_sym * p_f_a(Observed_state,Actual_state,2) *
P_state(Actual_state,current_treatment);

else

P_state_post(Resulting_state) = P_state_post(Resulting_state) +
P_asym_sym * p_f_a(Observed_state,Actual_state,2) *
P_state(Actual_state,current_treatment);
\begin{align*}
\text{P\_state\_post}(\text{Actual\_state}) &= \text{P\_state\_post}(\text{Actual\_state}) - \\
&\quad \text{P\_asym\_sym} \times \text{p\_f\_a}(\text{Observed\_state}, \text{Actual\_state}, 2) \times \\
&\quad \text{P\_state}(\text{Actual\_state}, \text{current\_treatment}); \\
&\text{end}
\end{align*}

\begin{align*}
\text{P\_state}(\text{Actual\_state}, \text{current\_treatment}); \\
&\text{end}
\end{align*}

\begin{align*}
\text{P\_state}(\text{current\_treatment}) &= \text{P\_state\_post};
\end{align*}

\begin{align*}
\% \text{FIT} \\
\text{if FIT\_AGE\_matrix}(\text{current\_treatment}, 2) == \text{age} \&\& \text{current\_treatment} < \\
((\text{max\_treatment}+1)/2)-1 \&\& \text{age} <= \text{stop\_age}
\text{P\_FIT}(\text{age\_index}) &= \text{P\_FIT}(\text{age\_index}) + \\
\text{sum}(\text{P\_state}(\text{current\_treatment}));
\end{align*}

\begin{align*}
\text{P\_negative} &= 0; \\
\text{P\_positive} &= 0; \\
\text{for i=1:max\_state}
\text{N\_small} &= 0; \\
\text{N\_medium} &= 0; \\
\text{N\_large} &= 0; \\
\text{N\_CRC} &= 0; \\
\text{for j=1:6}
\text{adenoma} &= \text{state\_list}(i,j+1);
\end{align*}
if adenoma==1
    N_small = N_small + 1;
elseif adenoma==2
    N_medium = N_medium + 1;
elseif adenoma==3
    N_large = N_large + 1;
elseif adenoma==4 || adenoma==5 || adenoma==6
    N_CRC = N_CRC + 1;
end
end

P_negative = P_negative + P_state(i,current_treatment) / 
sum(P_state(:,current_treatment)) * ((1-0.05)^N_small * (1-0.101)^N_medium * (1-
0.22)^N_large * (1-0.7)^(N_CRC));

P_positive = P_positive + P_state(i,current_treatment) / 
sum(P_state(:,current_treatment)) * (1 - ((1-0.05)^N_small * (1-0.101)^N_medium * (1-
0.22)^N_large * (1-0.7)^(N_CRC)));
end

negative_treatment = fit_matrix(current_treatment,2);
positive_treatment = fit_matrix(current_treatment,3);

P_state(:,negative_treatment) = P_negative * 
P_state(:,current_treatment);
P_state_local(:,:,negative_treatment) = P_negative * P_state_local(:,:,current_treatment);

P_state_regional(:,:,negative_treatment) = P_negative * P_state_regional(:,:,current_treatment);

P_state_distant(:,:,negative_treatment) = P_negative * P_state_distant(:,:,current_treatment);

P_state(:,positive_treatment) = P_positive * P_state(:,current_treatment);

P_state_local(:,:,positive_treatment) = P_positive * P_state_local(:,:,current_treatment);

P_state_regional(:,:,positive_treatment) = P_positive * P_state_regional(:,:,current_treatment);

P_state_distant(:,:,positive_treatment) = P_positive * P_state_distant(:,:,current_treatment);

P_state(:,current_treatment) = zeros(N_rows,1);

P_state_local(:,:,current_treatment) = zeros(N_rows,5,1);

P_state_regional(:,:,current_treatment) = zeros(N_rows,5,1);

P_state_distant(:,:,current_treatment) = zeros(N_rows,5,1);

% colonoscopy for positive FIT

P_state_post = P_state(:,positive_treatment);
for Actual_state=1:N_rows

\[ P_{\text{colonoscopy}}(\text{age_index}) = P_{\text{colonoscopy}}(\text{age_index}) + \]
\[ P_{\text{state}}(\text{Actual_state}, \text{positive_treatment}); \]

for Observed_state=1:N_rows

\[ \text{Group} = \text{state_group}($\text{Observed_state}$,2); \]
\[ \text{Resulting_state} = \text{colonoscopy}($\text{Actual_state}$,\text{Observed_state}); \]

if Resulting_state\(\neq 0\)

if Group==6

\[ P_{\text{state\_local}}(\text{Resulting_state},1,\text{positive_treatment}) = \]
\[ P_{\text{state\_local}}(\text{Resulting_state},1,\text{positive_treatment}) + \]
\[ p_f_a($\text{Observed_state}$,$\text{Actual_state}$,2) * P_{\text{state}}($\text{Actual_state}$,\text{positive_treatment}); \]
\[ P_{\text{state\_post}}(\text{Actual_state}) = P_{\text{state\_post}}(\text{Actual_state}) - \]
\[ p_f_a($\text{Observed_state}$,$\text{Actual_state}$,2) * P_{\text{state}}($\text{Actual_state}$,\text{positive_treatment}); \]

elseif Group==7

\[ P_{\text{state\_regional}}(\text{Resulting_state},1,\text{positive_treatment}) = \]
\[ P_{\text{state\_regional}}(\text{Resulting_state},1,\text{positive_treatment}) + \]
\[ p_f_a($\text{Observed_state}$,$\text{Actual_state}$,2) * P_{\text{state}}($\text{Actual_state}$,\text{positive_treatment}); \]
\[ P_{\text{state\_post}}(\text{Actual_state}) = P_{\text{state\_post}}(\text{Actual_state}) - \]
\[ p_f_a($\text{Observed_state}$,$\text{Actual_state}$,2) * P_{\text{state}}($\text{Actual_state}$,\text{positive_treatment}); \]

elseif Group==8

\[ P_{\text{state\_distant}}(\text{Resulting_state},1,\text{positive_treatment}) = \]
\[ P_{\text{state\_distant}}(\text{Resulting_state},1,\text{positive_treatment}) + \]
\[ p_f_a($\text{Observed_state}$,$\text{Actual_state}$,2) * P_{\text{state}}($\text{Actual_state}$,\text{positive_treatment}); \]
\[ P_{\text{state\_post}}(\text{Actual_state}) = P_{\text{state\_post}}(\text{Actual_state}) - \]
\[ p_f_a($\text{Observed_state}$,$\text{Actual_state}$,2) * P_{\text{state}}($\text{Actual_state}$,\text{positive_treatment}); \]
\[ P_{\text{state\_post}}(\text{Actual\_state}) = P_{\text{state\_post}}(\text{Actual\_state}) - p_{f\_a}(\text{Observed\_state},\text{Actual\_state},2) \times P_{\text{state}}(\text{Actual\_state},\text{positive\_treatment}); \]

\text{else}

\[ P_{\text{state\_post}}(\text{Resulting\_state}) = P_{\text{state\_post}}(\text{Resulting\_state}) + p_{f\_a}(\text{Observed\_state},\text{Actual\_state},2) \times P_{\text{state}}(\text{Actual\_state},\text{positive\_treatment}); \]

\[ P_{\text{state\_post}}(\text{Actual\_state}) = P_{\text{state\_post}}(\text{Actual\_state}) - p_{f\_a}(\text{Observed\_state},\text{Actual\_state},2) \times P_{\text{state}}(\text{Actual\_state},\text{positive\_treatment}); \]

\text{end}

\text{end}

\text{end}

\text{end}

\[ P_{\text{state}}(:,\text{positive\_treatment}) = P_{\text{state\_post}}; \]

\text{end}

\text{end}

\text{end}

\% yearly probabilities

\[ P_{\text{Asymptomatic}}(\text{age\_index}) = \text{sum(sum}(P_{\text{state}})); \]

\text{for} i=1:5

\[ P_{\text{local}}(\text{age\_index},i) = \text{sum(sum}(P_{\text{state\_local}}(:,i,:))); \]

\[ P_{\text{regional}}(\text{age\_index},i) = \text{sum(sum}(P_{\text{state\_regional}}(:,i,:))); \]

\[ P_{\text{distant}}(\text{age\_index},i) = \text{sum(sum}(P_{\text{state\_distant}}(:,i,:))); \]

\text{end}

\text{age}
\[ 100 \times (P\text{\_death}(\text{age\_index}) + P\text{\_Asymptomatic}(\text{age\_index}) + \\
\text{sum}(P\text{\_local}(\text{age\_index,:,})) + \text{sum}(P\text{\_regional}(\text{age\_index,:,})) + \\
\text{sum}(P\text{\_distant}(\text{age\_index,:,}))) \\
100 \times [P\text{\_death}(\text{age\_index}); P\text{\_Asymptomatic}(\text{age\_index}); \\
\text{sum}(P\text{\_local}(\text{age\_index,:,})); \text{sum}(P\text{\_regional}(\text{age\_index,:,})); \text{sum}(P\text{\_distant}(\text{age\_index,:,})); \\
P\text{\_colonoscopy}(\text{age\_index}); P\text{\_FIT}(\text{age\_index})] \\
\text{end} \\
\text{end} \]
30. TOTAL_COSTS_FIT_discounting

This function calculates the total costs at each time step.

TOTAL_COSTS_FIT_discounting(P_local, P_regional, P_distant, P_colonoscopy, P_FIT, T1)

function [total_costs] = TOTAL_COSTS_FIT_discounting(P_local, P_regional, P_distant, P_colonoscopy, P_FIT, T1)

global Initial_Local_CRC_costs Surveillance_Local_CRC_costs
Initial_Regional_CRC_costs Surveillance_Regional_CRC_costs
Initial_Distant_CRC_costs Surveillance_Distant_CRC_costs Terminal_CRC_costs
Cost_of_colonoscopy Cost_Adverse_effects Cost_of_FIT

global dt age_max total_ages age_min

global P_local_death P_regional_death P_distant_death

t_0 = T1;

r = 0.03;

Local_CRC_costs=zeros(total_ages,1);
Regional_CRC_costs=zeros(total_ages,1);
Distant_CRC_costs=zeros(total_ages,1);
Colonoscopy_Costs=zeros(total_ages,1);
FIT_Costs=zeros(total_ages,1);
Z=zeros(total_ages,1);
for age=age_min:dt:age_max

\[ T = (\text{age} - \text{age}_{\text{min}})/\text{dt} + 1; \]

\[ t_1 = \text{age} - t_0; \]

if age\(<t_0$

\[ Z(T) = 0; \]

else

\[ Z(T) = r \times (1/(r \times \exp(r \times 0)) - 1/(r \times \exp(r \times t_1))); \]

end

Colonoscopy\_Costs(T) = (\text{Cost\_of\_colonoscopy} + \text{Cost\_Adverse\_effects}) \times \]

\[ P_{\text{colonoscopy}}(T); \]

FIT\_Costs(T) = Cost\_of\_FIT \times P_{\text{FIT}}(T);

Local\_CRC\_costs(T)= Initial\_Local\_CRC\_costs \times P_{\text{local}}(T,1);

Local\_CRC\_costs(T)=Local\_CRC\_costs(T)+Surveillance\_Local\_CRC\_costs \times P_{\text{local}}(T,2);

Local\_CRC\_costs(T)=Local\_CRC\_costs(T)+Surveillance\_Local\_CRC\_costs \times P_{\text{local}}(T,3);

Local\_CRC\_costs(T)=Local\_CRC\_costs(T)+Surveillance\_Local\_CRC\_costs \times P_{\text{local}}(T,4);
\[ \text{Local\_CRC\_costs}(T) = \text{Local\_CRC\_costs}(T) + \text{Surveillance\_Local\_CRC\_costs} \times P_{\text{local}}(T,5); \]
\[ \text{Local\_CRC\_costs}(T) = \text{Local\_CRC\_costs}(T) + \text{Terminal\_CRC\_costs} \times P_{\text{local}}(T,1) \times P_{\text{local\_death}}(1); \]
\[ \text{Local\_CRC\_costs}(T) = \text{Local\_CRC\_costs}(T) + \text{Terminal\_CRC\_costs} \times P_{\text{local}}(T,2) \times P_{\text{local\_death}}(2); \]
\[ \text{Local\_CRC\_costs}(T) = \text{Local\_CRC\_costs}(T) + \text{Terminal\_CRC\_costs} \times P_{\text{local}}(T,3) \times P_{\text{local\_death}}(3); \]
\[ \text{Local\_CRC\_costs}(T) = \text{Local\_CRC\_costs}(T) + \text{Terminal\_CRC\_costs} \times P_{\text{local}}(T,4) \times P_{\text{local\_death}}(4); \]
\[ \text{Local\_CRC\_costs}(T) = \text{Local\_CRC\_costs}(T) + \text{Terminal\_CRC\_costs} \times P_{\text{local}}(T,5) \times P_{\text{local\_death}}(5); \]

\[ \text{Regional\_CRC\_costs}(T) = (\text{Initial\_Regional\_CRC\_costs}) \times P_{\text{regional}}(T,1); \]
\[ \text{Regional\_CRC\_costs}(T) = \text{Regional\_CRC\_costs}(T) + \text{Surveillance\_Regional\_CRC\_costs} \times P_{\text{regional}}(T,2); \]
\[ \text{Regional\_CRC\_costs}(T) = \text{Regional\_CRC\_costs}(T) + \text{Surveillance\_Regional\_CRC\_costs} \times P_{\text{regional}}(T,3); \]
\[ \text{Regional\_CRC\_costs}(T) = \text{Regional\_CRC\_costs}(T) + \text{Surveillance\_Regional\_CRC\_costs} \times P_{\text{regional}}(T,4); \]
\[ \text{Regional\_CRC\_costs}(T) = \text{Regional\_CRC\_costs}(T) + \text{Surveillance\_Regional\_CRC\_costs} \times P_{\text{regional}}(T,5); \]
Regional_CRC_costs(T)=Regional_CRC_costs(T)+Terminal_CRC_costs * P_regional(T,1)*P_regional_death(1);

Regional_CRC_costs(T)=Regional_CRC_costs(T)+Terminal_CRC_costs * P_regional(T,2)*P_regional_death(2);

Regional_CRC_costs(T)=Regional_CRC_costs(T)+Terminal_CRC_costs * P_regional(T,3)*P_regional_death(3);

Regional_CRC_costs(T)=Regional_CRC_costs(T)+Terminal_CRC_costs * P_regional(T,4)*P_regional_death(4);

Regional_CRC_costs(T)=Regional_CRC_costs(T)+Terminal_CRC_costs * P_regional(T,5)*P_regional_death(5);

Distant_CRC_costs(T)= (Initial_Distant_CRC_costs) * P_distant(T,1);

Distant_CRC_costs(T)=Distant_CRC_costs(T)+Surveillance_Distant_CRC_costs * P_distant(T,2);

Distant_CRC_costs(T)=Distant_CRC_costs(T)+Surveillance_Distant_CRC_costs * P_distant(T,3);

Distant_CRC_costs(T)=Distant_CRC_costs(T)+Surveillance_Distant_CRC_costs * P_distant(T,4);

Distant_CRC_costs(T)=Distant_CRC_costs(T)+Surveillance_Distant_CRC_costs * P_distant(T,5);

Distant_CRC_costs(T)=Distant_CRC_costs(T)+Terminal_CRC_costs * P_distant(T,1)*P_distant_death(1);
\begin{align*}
\text{Colonoscopy\_Costs} &= \text{Colonoscopy\_Costs} - (\text{Colonoscopy\_Costs} \times Z); \\
\text{FIT\_Costs} &= \text{FIT\_Costs} - (\text{FIT\_Costs} \times Z); \\
\text{Local\_CRC\_costs} &= \text{Local\_CRC\_costs} - (\text{Local\_CRC\_costs} \times Z); \\
\text{Regional\_CRC\_costs} &= \text{Regional\_CRC\_costs} - (\text{Regional\_CRC\_costs} \times Z); \\
\text{Distant\_CRC\_costs} &= \text{Distant\_CRC\_costs} - (\text{Distant\_CRC\_costs} \times Z); \\
\text{total\_costs} &= \text{sum}(\text{Colonoscopy\_Costs} + \text{Local\_CRC\_costs} + \\
&\quad \text{Regional\_CRC\_costs} + \text{Distant\_CRC\_costs} + \text{FIT\_Costs});
\end{align*}
31. TOTAL_QALYs_discounting_FIT

This function calculates the total QALYs at each time step.

TOTAL_QALYs_discounting_FIT (P_local, P_regional, P_distant, P_death, P_Asymptomatic, T1)

function [total_qalys] = TOTAL_QALYs_discounting_FIT (P_local, P_regional, P_distant, P_death, P_Asymptomatic, T1)
global QALYs_healthy QALYs_dead QALYs_Local_CRC
QALYs_Regional_CRC QALYs_Distant_CRC

global dt age_max total_ages age_min

t_0 = T1;

r=0.03;

QALYs = zeros(total_ages,1);
Z=zeros(total_ages,1);

for age=age_min:dt:age_max
    T = (age - age_min) / dt + 1;
    t_1 = age-t_0;
    if age<=t_0
        Z(T) = 0;
    else
        Z(T) = r * (1/(r*exp(r*0)) - 1/(r*exp(r*t_1)));
    end
    QALYs(T) = QALYs(T) + QALYs_healthy * P_Asymptomatic(T);
    QALYs(T) = QALYs(T) + QALYs_dead * P_death(T);
    QALYs(T) = QALYs(T) + QALYs_Local_CRC * sum(P_local(T,:));
    QALYs(T) = QALYs(T) + QALYs_Regional_CRC * sum(P_regional(T,:));
QALYs(T) = QALYs(T) + QALYs_Distant_CRC * sum(P_distant(T,:));

end

QALYs = QALYs-(QALYs.*Z);

total_qalys = sum(QALYs);

end

32. FIT_MATRIX ()

This function creates a matrix to know when a positive interval and a negative interval for a FIT are taken for each treatment.

function [fit_matrix] = FIT_MATRIX ()

[treatments, treatment_list, max_treatment] = TREATMENT_DEFINITIONS_7();

fit_matrix = zeros(127,3);
for current_treatment=1:127

i = 0;

while treatment_list(current_treatment,i+1)~=0
    i = i + 1;
end

for j=1:7
    if j==i
        negative_treatment(i) = 1;
        positive_treatment(i) = 3;
    else
        negative_treatment(j) = treatment_list(current_treatment,j+1) + 2;
        positive_treatment(j) = treatment_list(current_treatment,j+1) + 2;
    end
end

negative_number = treatments(negative_treatment(1), negative_treatment(2),
negative_treatment(3), negative_treatment(4), negative_treatment(5),
negative_treatment(6), negative_treatment(7));

positive_number = treatments(positive_treatment(1), positive_treatment(2),
positive_treatment(3), positive_treatment(4), positive_treatment(5),
positive_treatment(6), positive_treatment(7));
fit_matrix(current_treatment,:) = [current_treatment, negative_number, positive_number];

end

33. FIT_AGE (start_age, positive_step, negative_step)

This function gives you all the combinations for the age of each positive and negative interval.

function [FIT_AGE_matrix] = FIT_AGE (start_age, positive_step, negative_step)

[treatments, treatment_list, max_treatment] = TREATMENT_DEFINITIONS_7();
FIT_AGE_matrix=zeros(255,2);
for treatment = 1:255
    FIT_AGE = start_age;
    for j= 2:6
        if treatment_list(treatment,j) == 1 ;
            FIT_AGE = FIT_AGE + positive_step;
        elseif treatment_list( treatment,j) == -1;
            FIT_AGE = FIT_AGE + negative_step;
        end
    end
end
FIT_AGE_matrix(treatment, 1) = treatment;
FIT_AGE_matrix(treatment, 2) = FIT_AGE;
BIBLIOGRAPHY


http://www.indiahealthhelp.com/procedure/colon-cancer-treatment/


