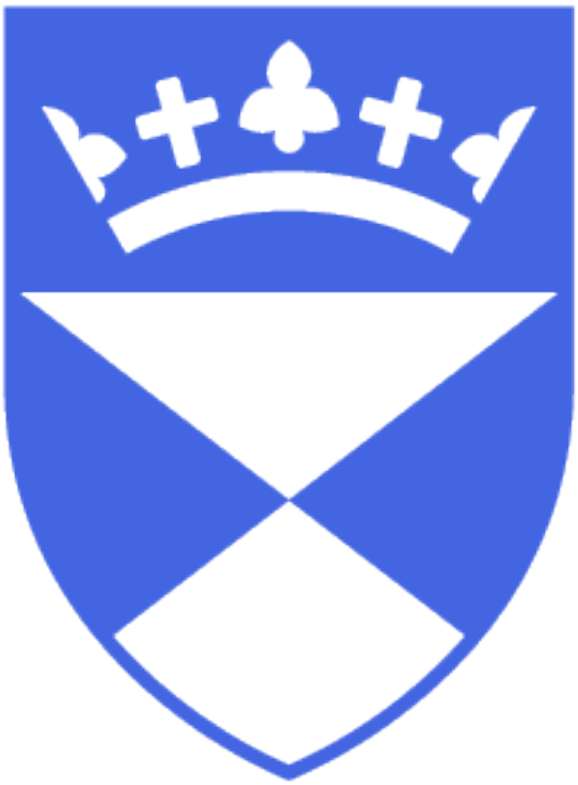
MASTER OF SCIENCE

Assessing the effectiveness of current UK guidelines on familial colorectal cancer risk

Christou, Koula

Award date:
2020

Assessing the effectiveness of current UK guidelines on familial colorectal cancer risk



Kyriaki Christou

MSc by Research in Medicine
School of Medicine
University of Dundee
July 2020

Table of Contents

Acknowledgements

I would firstly like to thank my supervisor, Dr Jonathan Berg for the invaluable guidance and insightful comments throughout this project. I am very thankful for always being available for discussion and for enriching my knowledge in cancer genetics through the importance of research.

I would also like to thank Dr David Goudie and Mrs Jacqueline Dunlop from the clinical genetics department for assisting with data availability and queries at various stages of this research.

My sincerest gratitude also goes to Dr James Cotton and Dr Richard Oparka for being so willing to help and share information from their departments' databases for the completion of this project.

A special thank you to Dr Ellie Hothersall for her support in both my masters and medicine degree.

I also wish to acknowledge Dr James Abbott and Mr Marek Gierliniski from the School of Life Sciences for their statistical input.

To my dear friend Cleanthis Michael, for his statistical advice, proofreading skills, patience and support, I am forever grateful.

I would also like to express my appreciation to my friends and colleagues Thineskrishna Anbarasan and Pavithan Maniam for their help with data transferring and processing.

Lastly, my thanks to my loving family and my partner Spyridonas Georgiou for their daily encouragement and care throughout the process of completing this work.

Declaration

I declare that the content of this project report is my own work and has not previously been submitted for any other assessment. The report is written in my own words and conforms to the University of Dundee's Policy on plagiarism and academic dishonesty. Unless otherwise indicated, I have consulted all the references cited in this report:

Signed: 

Date: 6th July 2020

Summary:

Introduction: Family history (FH) of colorectal cancer (CRC) is a frequent reason for referral to Clinical Genetics in the UK. The British Society of Gastroenterologists (BSG) guideline stratifies patients to risk categories (low/population, low-moderate, high-moderate and high) according to FH. Individuals with Lynch syndrome are classified differently to those who have a high-risk FH, but no high penetrance mutation. We investigated how effectively BSG guidelines categorise people at increased risk of CRC.

Methods: FH data was obtained for all unaffected people with a family history of CRC, referred to Tayside clinical genetics from 2000-2009. Risk category according to BSG guidance was assigned de novo. Individuals who went on to develop adenomatous polyps or CRC were identified by record linkage.

Results: 1120 patients were identified and after exclusion criteria, there were 728 non-polyposis patients (288 low-risk, 316 moderate-risk and 121 high-risk, including 31 mutation carriers). 8 invasive CRC developed, 2 in low, 3 in moderate and 3 in high-risk groups. There was no significant difference in the Relative Risk (RR) of cancer development between groups. The only significant finding was an increased risk of CRC in mutation carriers, RR 9.290 (1.3557-63.6653). Kaplan-Meier analysis demonstrated no significant difference in cancer rates between groups. There was a significantly higher risk of polyp detection in the high-risk group compared to the low-risk group. Kaplan-Meier analysis demonstrated a significantly higher likelihood of polyp detection in the high-risk group when compared to both low and moderate-risk groups.

Conclusions: Presence of mutation seems to be the best predictor of cancer risk. Colonoscopic surveillance may be effective in reducing the cancer incidence in the moderate and high-risk groups. The study re-affirms that no colonoscopic screening is required in the moderate-risk group aged less than 50. Furthermore, it may suggest that less screening is required in the high-risk group beyond the age of 50.

List of Abbreviations:

AR: Absolute Risk;
APC: Adenomatous Polyposis Coli;
AJCC: America Joint Committee of Cancer;
ACMG: American College of Medical Genetics and Genomics;
ANOVA: Analysis of Variance;
AFAP: Attenuated Familial Adenomatous Polyposis;
CRC: Colorectal Cancer;
CHI: Community Health Index;
CHRPE: Congenital Hypertrophy of Retinal Pigment Epithelium;
ESMO: European Society for Medical Oncology;
FOBt: Faecal Occult Blood test;
FAP: Familial Adenomatous Polyposis;
FH: Family History;
FDR: First Degree Relative;
GWAS: Genome-Wide Association Studies;
HHT: Hereditary Haemorrhagic Telangiectasia;
HNPCC: Hereditary Non-Polyposis Colorectal Cancer;
IARC: International Agency for Research on Cancer;
IHC: Immunohistochemistry;
IBD: Inflammatory Bowel Disease;
JPS: Juvenile Polyposis Syndrome;
KM: Kaplan Meier;
MSI: Microsatellite Instability;
MMR: Mismatch Repair;
MAP: MUTYH-Associated Polyposis;
NCCN: National Comprehensive Cancer Network;
NICE: National Institute of Health and Care Excellence;
NPV: Negative Predictive Value;
OR: Odds Ratio;
OGD: Oesophago-Gastro-Duodenoscopy;
PJS: Peutz-Jeghers Syndrome;
PRS: Polygenic Risk Scores;
PPV: Positive Predictive Value;
PREMM: PRediction Model for Gene Mutations;
FIT: Faecal Immunochemical Test;
RCT: Randomised Controlled Trial;
RR: Relative Risk;
SIGN: Scottish Intercollegiate Guidelines Network;
SDR: Second Degree Relative;
SNPs: Single Nucleotide Polymorphisms;
BSG: The British Society of Gastroenterologists,
TNM: Tumour-Node-Metastases;
VUS: Variant of Uncertain Significance;
VSE: Video Capsule Endoscopy.

List of Figures:

List of Tables:

Chapter 1: Introduction

1.1 Overview of colorectal cancer

1.1.1 Epidemiology

According to the International Agency for Research on Cancer data in 2018, colorectal cancer (CRC) ranks as the third most common cancer worldwide, with 1.8 million new diagnoses.(1) The incidence patterns are reflected equally in both males and females, however, there is geographical variation with the highest incidence in Australia and New Zealand and the lowest incidence in Western Africa. When comparing more with less developed regions, 55% of the cases occur in the former.(1,2) In the UK, CRC accounts for 12% of all new cancer cases, with 55% of those occurring in males. The rates have remained stable over the last decade; it is estimated that 1 in 14 men and 1 in 19 women will develop CRC in their lifetime.(3) The incidence is strongly correlated with age, rising steeply from around the age of 50 (*See* **Figure 1**).(3) Mortality rates also vary worldwide, with fewer deaths in more developed regions, reflecting better survival due to better treatment.(1) Even though mortality rates have decreased by 42% in the last 50 years, CRC remains the second most common cause of cancer death in the UK. Mortality also increases with age, with higher rates in males above the age of 60 years old (*See* **Figure 2**).(3–6)

In Scotland, the incidence of CRC for the period of 2013-2017 was 15,127 cases. (7) Mortality rate was 31.1%, which is higher than England (24.4%) in 2017.(6,7) Age specific incidence in Scotland follows the same trend as the rest of the UK; however, the age-specific mortality rate is higher between ages 70-74, rather than 80-84 as shown in **Figure 2**.

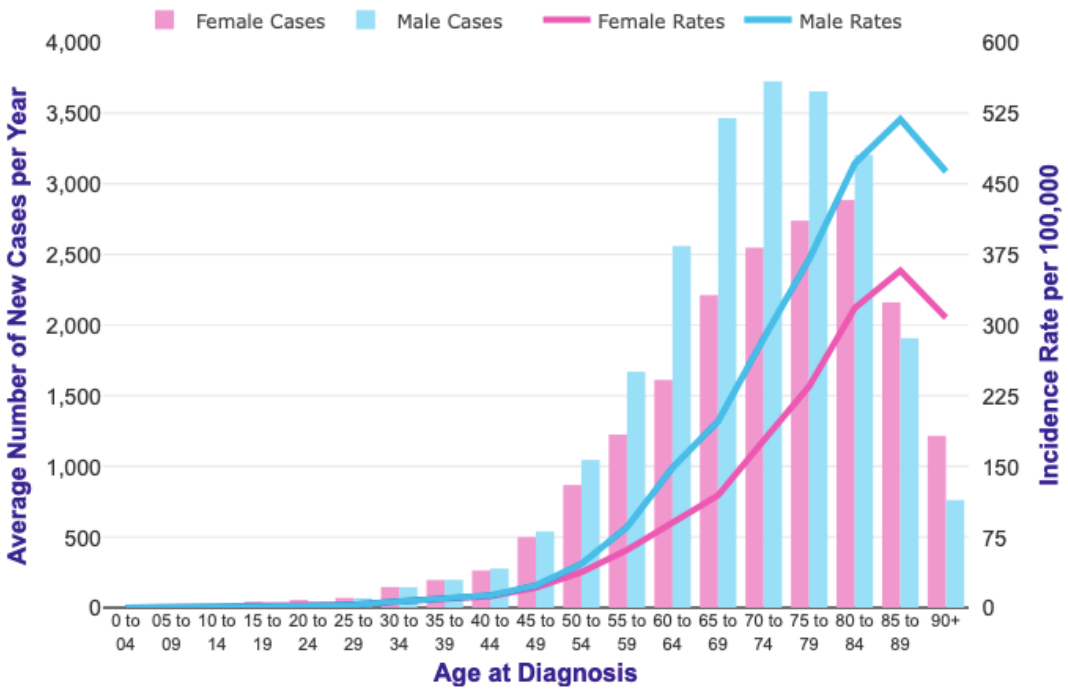


Figure 1 *Incidence of CRC by age and gender.* Figure from source (3).

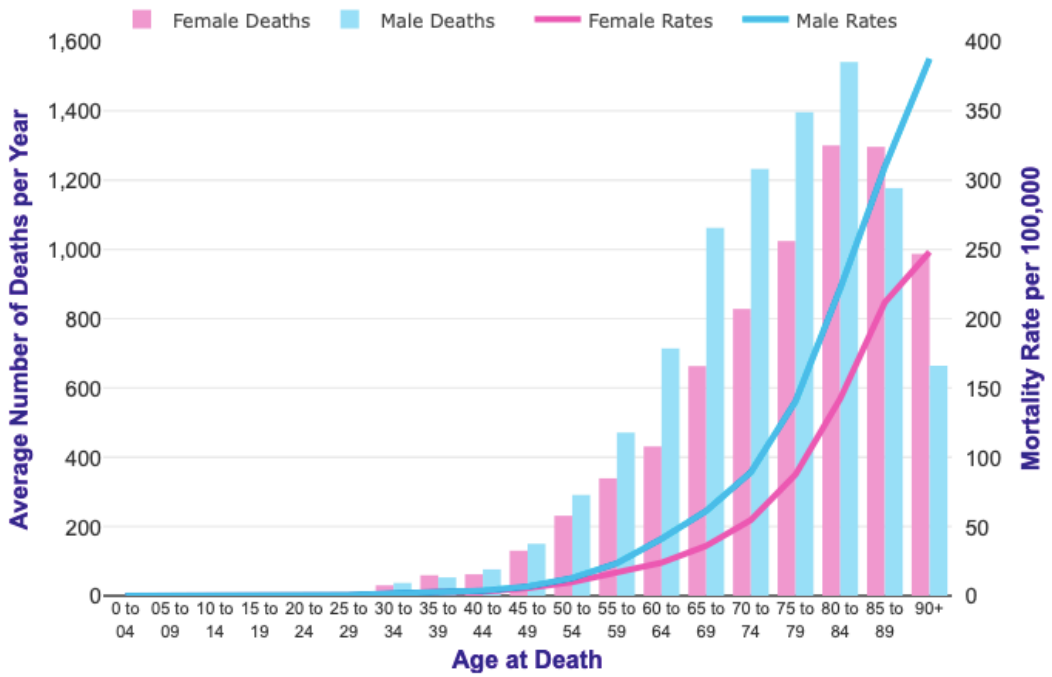


Figure 2 *CRC mortality by age and gender.* Figure from source 7.

1.1.2 Population Screening:

Due to the increased incidence of CRC, as well as the fact that most cancers were identified at a later stage (*See* **Section 1.2**), a method needed to be implemented to detect CRC at an early, curable stage. Thus, a screening programme, which aims to detect non-visible blood in the faeces- faecal occult blood test (FOBt) was implemented in various countries. A variety of high-quality trials were conducted to assess the effectiveness of this test. One of the first randomised controlled trials (RCT) included 45000 people aged 50-80 years who were randomised to receive no FOBt screening (control group) or receive FOBt screening every two years. The cumulative incidence of CRC development was significantly reduced in the intervention groups (relative risk 0.80, 95% CI 0.70 to 0.90; p= 0.002).(8) Moreover, three RCTs conducted in the US, UK and Denmark revealed reduced mortality of CRC in people who undertook screening.(8–10) This was also confirmed in a Cochrane review and meta-analysis of these studies, revealing a reduction in mortality by 16% in the screened population (relative risk 0.84, 95% CI 0.78 to 0.90).(11)

In the UK, bowel screening was implemented after pilot studies were conducted in England and Scotland.(12) The studies' outcomes were in agreement with the previous RCTs, demonstrating a significant decrease in bowel cancer mortality. In England, Wales and Northern Ireland, screening is offered every two years between the ages of 60-74. However, people in England and Scotland that exceeded 74 years of age may be screened upon request whereas the rest of the countries do not accept requests. In Scotland, screening is implemented from 50-74 years of age.(13) Within 6 months of invitation, it is estimated that around 50% of the people who are invited across the UK are screened with a definitive result. Screening uptake is higher in females than in males across the whole of the UK.(14) Screening uptake data are available for England, showing a reduced uptake in the younger age groups (53.5% in 60-64 age group compared to the 62% and 60.3% in the 65-69 and 70-74 age groups).(14)

If the initial test is not definitive (unclear), two repeat FOBts are sent in England or a faecal immunochemical test (FIT) in Wales, Scotland and Northern

Ireland.(3) People who have an abnormal result will have a colonoscopy, unless this method is deemed inappropriate or the person does not attend their appointment. People will have different management according to what is detected at the time of colonoscopy. If no abnormality is detected, the FOBt will be offered again every two years.

Initial FOBt has a higher positivity rate because true positives are taken out of the population for subsequent (incident) screens. Following colonoscopy, adenomas are found in 48% of men and 35% of women in England, with a high proportion of those being intermediate or high risk. Bowel cancer is found in 12-15% of men and 8% of women who were investigated after an abnormal screening result.(14) After the initial screen, people with abnormal incident screens have about 5% chance of being diagnosed with CRC and 25% chance of having a benign polyp. Even though a number of people with CRC have been identified through FOBt, no study has compared FOBt to colonoscopy in order to determine the specificity or negative predictive value of the test in asymptomatic population. Approximately 60% of people with CRC will have abnormal FOBt result.(13,15,16)

Since 2017, Quantitative faecal immunochemical testing (FIT) has been implemented as first-line screening method in the Scottish Bowel Screening programme, with the screening age group and the recall time remain the same. Studies have shown an increased uptake of the FIT test across gender, age and deprivation categories, which may be attributed to its simpler use and to the fact that it only requires one sample to be taken.(17,18) This test measures micrograms of human haemoglobin per gram of faeces (µg Hb/g faeces) rather than just the presence of blood; a result of more than 80 µg Hb/g faeces will be referred for further assessment with colonoscopy. It is also specific to human haemoglobin and is less likely to be affected by diet and certain medications. The positivity is higher when using FIT rather than FOBt (3.1% and 1.9% respectively). Moreover, FIT is a better positive predictor in identifying adenomas than FOBt (43.5% vs 40.00%). Even though FOBt is more accurate in identifying CRC, cancer detection is higher using FIT due to increased uptake and greater percentage of positive tests. Identified CRCs through colonoscopy after positive

FIT test were mostly (60.3%) in the first two stages of disease with no metastasis.(19,20)

Even though the initial testing has no direct associated risks, it is possible that people might develop anxiety or have false reassurance due to a negative screening test. Furthermore, colonoscopy has a separate set of risks and complications, such as heavy bleeding, bowel perforation and even death. People in high risk categories such as those with previous diagnosis of CRC, colonic adenomas, inflammatory bowel disease and acromegaly have separate guidelines.(13) Different guidelines also apply to people with strong family history or genetic predisposition to CRC development (*See* **Section 1.7**).

When considering the cost-effectiveness of the screening programme, in a meta-analysis of 55 publications, all studies concluded that any form of CRC screening is more cost-effective or even cost-saving compared to no screening.(21) Furthermore, several studies have reported the superiority of FOBt either alone or in combination with other screening methods such as sigmoidoscopy in terms of cost-effectiveness for cancer mortality.(22–24) A recent study comparing the cost-effectiveness of FOBt with FIT, suggests that FIT is significantly ($p<0.001$) cost-saving and results in quality adjusted life years gains of 0.014 (95% CI 0.012 to 0.017). However, due to the pressures on endoscopy services, alternative screening programmes are not feasible.(24,25) Furthermore, the impact of the screening programme on hospital diagnostic services must also be considered. Nevertheless, the benefits including cost-effectiveness as well as incidence and mortality outweigh the harms.(26)

1.2 Pathology

1.2.1 Polyps

Polyps can be defined as small growths that can occur throughout the GI tract, most commonly in the colon. Polyps can either grow without stalks, called sessile, or protrude from the mucosa surface, termed pedunculated. Generally, polyps can be classified in non-neoplastic and neoplastic. There are three types of non-

neoplastic polyps, the first one known as hyperplastic which is thought to arise due to decreased epithelial turnover and devalued shedding, resulting in a “pile-up” of goblet cells. The second, inflammatory polyp is part of a solitary rectal ulcer syndrome and patients can present with rectal bleeding, mucus discharge and inflammation in the anterior rectal wall. The third type of non-neoplastic polyp is called hamartomatous and can occur either sporadically or as components of some inherited syndromes **(***See* **Section 1.5**). The most common type of neoplastic polyps are adenomas, which are infrequent in African and Asian countries but occur in nearly 50% of the Western population aged over 50. Adenomas can be pedunculated or sessile and can range from 0.3 to 10cm in diameter. They are characterised by the presence of epithelial dysplasia, which can be identified histologically as nuclear hyperchromasia, elongation and stratification (*See* **Figure 3**). Adenomatous polyps can be sub-classified into tubular, tubulovillous or villous on the basis of their architecture; however, evidence suggests little clinical significance of this differentiation. Sessile serrated adenomas are most commonly found in the right colon, lack dysplasia and their malignant potential is similar to the typical adenomas.(27,28)

Even though adenomas are benign lesions, evidence from epidemiological studies suggests that they are precursors of CRC. Studies show that the incidence by geographic location of CRC follows the adenoma pattern and that the prevalence of adenomas peaks at least 5 years prior to CRC development. Furthermore, malignant foci have been identified in adenomas as well as remnants of adenomas in CRC.(29–32) Removal of adenomas can decrease the incidence of CRC by up to 90% as well as mortality, further supporting that adenomas might be precursors of CRC.(33–35) The rate of transformation of adenoma to adenocarcinoma is around 0.25%.(36) There is a strong positive correlation between size of adenoma and risk of malignancy, with studies suggesting that around 40% of lesions larger than 4cm contain malignant foci. (27,29)

The British Society of Gastroenterologists (BSG) have guidance on surveillance following adenoma detection.(37) The frequency of surveillance depends on the patient’s risk category, according to the number and size of adenomatous polyps

found at baseline colonoscopy (*See* **Table 1** and **Figure 4).** During the endoscopy procedure, it is recommended that scanning of the colonic mucosa occurs both during insertion and withdrawal of the colonoscope to reduce the miss rate of small polyps. Pancolonic dye spraying is also used to aid in the detection of small flat polyps. It is also pivotal to completely excise the polyp, as studies have shown that a substantial percentage of CRCs develop at the site of previous polypectomy.(37–39) In inherited polyposis conditions, different guidelines apply (*See* **Section 1.7**).

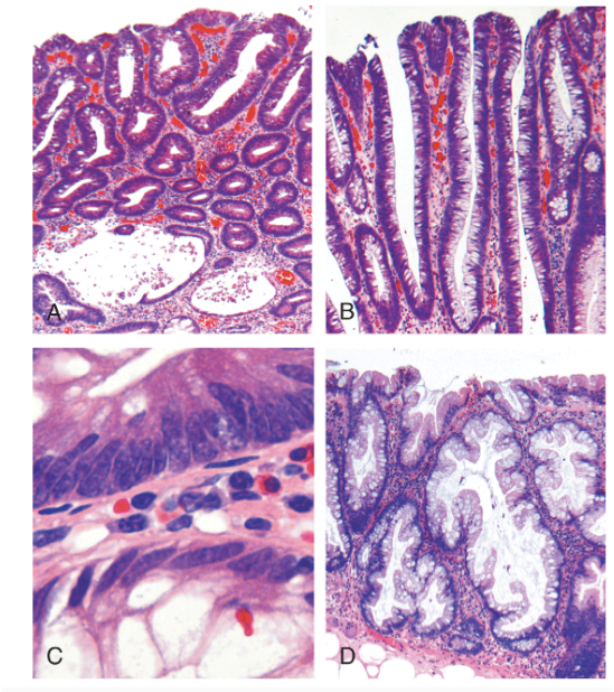


Figure 3 *Histology of adenomatous polyps in the colon. A. Tubular adenoma; B. Villous adenoma; C. Dysplastic epithelial cells; D. Sessile serrated adenoma. Picture taken from source (27).*

Table 1 *BSG risk criteria for adenomas.* Information taken from source 27.

BSG Risk Category	*Colonoscopy findings*
High risk	≥5 small adenomas OR ≥3 at least one ≥1 cm
Intermediate risk	3-4 small adenomas OR At least one ≥1cm
Low risk	1-2 adenomas AND Both small <1cm

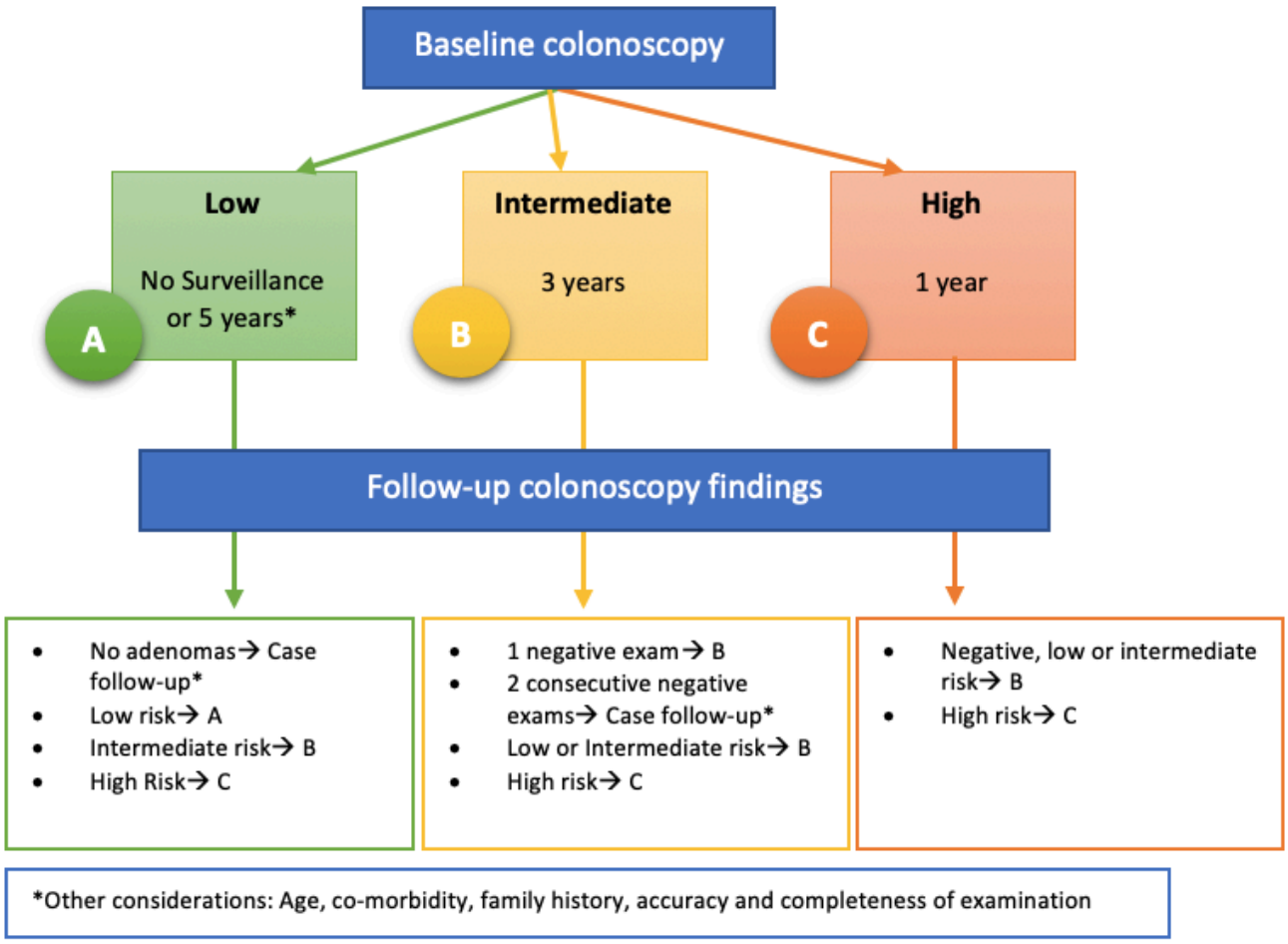


Figure 4 *Recommended surveillance according to colonoscopic findings.* Information extracted from source *37.*

1.2.2 The adenoma-carcinoma sequence

Even though not all adenomas become malignant, an estimated of 85% of CRC are thought to have an adenoma precursor. There are various mechanisms involved in the pathogenesis of CRC and how carcinomas evolve from adenomas.(27,40,41) Both genetic and epigenetic abnormalities and molecular events are responsible for the adenocarcinoma formation. One of the most well-established pathways is that of APC/β-catenin, which accounts for around 80% of sporadic CRC. APC is a tumour suppressor gene and a negative regulator of β-catenin. Loss of APC function leads to failure of degradation of β-catenin, leading to accumulation in the nucleus and increased transcription of genes such as MYC and Cyclin D1, involved in proliferation. The increased proliferation is followed by prevention of apoptosis due to activating mutations of the KRAS gene, which is seen in 50% of adenomas that are greater than 1 cm in diameter and in 50% of invasive CRC. Defects in DNA repair pathways can contribute to the accumulation of somatic mutations that lead to uncontrolled proliferation or enhanced survival of neoplastic cells. Constitutional and somatic mutations in DNA mismatch repair genes can cause instability of polyA tracts in the Epidermal Growth factor receptor gene (42) and other simple sequence repeats in genes involves in growth regulation such as TGFBR2 and BAX. Furthermore, complex mechanisms including the CpG island methylator phenotype (CIMP) have been implicated.(27,41,43,44) The mechanisms in which carcinogenesis occurs can overlap and are used in the molecular classification, as well as prognosis and management.(27)

1.2.3 Colorectal Cancer Staging

CRC tumours are typically classified by the criteria of the America joint committee of cancer (AJCC) with the Tumour-Node-Metastases (TNM) system.(45) Usually, T1 and T2 have tumours just invading through the submucosa but not into the muscular propria; their 5-year survival rate is 75%.(46) However, stage IV has very poor prognosis, with 6% survival rate.(46) Duke's staging is also widely used to classify CRC (*See* **Table 2**).(47) Treatment depends on staging, always taking into account the person's fitness depending on age and other comorbidities as well as their wishes. Microsatellite instability (MSI) testing of the tumour is also offered in people with CRC as it predicts treatment response. Depending on the type of treatment, intensive surveillance with colonoscopy may still be required, based on the BSG recommendations.(37)

Table 2 *Dukes staging of colorectal carcinoma and 5-year survival. (46,47)*

Stage	**Features**	**5-year survival**
A	Tumour confided in the mucosa	95-100%
B1	Tumour growth into muscularis propria	80-90%
B2	Tumour growth through muscularis propria and serosa (full thickness)	80-90%
C1	Tumour spread to 1-4 regional lymph nodes	65%
C2	Tumour spread to more than 4 regional lymph nodes	
D	Distant metastasis (liver, lung, bones)	5-10%

1.3 Modifiable lifestyle factors and CRC risk

A number of modifiable lifestyle factors have been implicated to be related to CRC. It is estimated that 54% of CRC cases could have been prevented with lifestyle change.(48,49)

Table 3 shows the preventability estimates of CRC in the UK for each lifestyle factor. **Table 4** summarises the lifestyle factors that protect and increase risk of colorectal cancer.

Table 3 *Approximate percentage of preventable CRC cases in the UK according to lifestyle factors. (48,49)*

Lifestyle factor	Approximate percentage of CRC cases preventable in the UK (%)
Insufficient fibre consumption	28
Processed and red meat	13
Body fatness and obesity	11

Table 4 *Lifestyle protective and risk factors for colorectal cancer.*

Protective factors	Risk factors
Wholegrains	Red & processed meat
Dietary fibre	Alcoholic drinks
Weight loss	Body fatness and weight increase
	Smoking

1.3.1 Diet and nutrition

1.3.1.1 Red and processed meat

Red meat is classified as a probable cause of colorectal cancer. Red meat contains high levels of haem iron, which has been proven to simulate endogenous formation of carcinogenic compounds, promoting CRC tumorigenesis. Furthermore, due to the high temperature of cooking, heterocyclic amines and polycyclic aromatic hydrocarbons are formed, which promote carcinogenesis.(50–52)

The definition of “processed” generally describes meat that has been transformed to enhance flavour or preservation, through salting, curing, smoking and other processes. Types of processed meat include ham, bacon, salami, pastrami and some sausages.(50) Consumption of 50 grams of processed meat per day significantly increases the risk of CRC development by 16% (RR 1.16, 95% CI 1.08-1.26), classifying it as a convincing cause. Processed meat can contribute to carcinogenesis in a similar mechanism to red meat. Moreover, the high fat content is a source of N-nitroso compounds and can also be responsible for the production of secondary bile acids, leading to tumorigenesis.(50,51,53)

1.3.1.2 Wholegrains and dietary fibre

Wholegrains or cereals are a category of energy stores of grain seed; the main types include heat, rice, barley, oats and rye. Wholegrain intake seems to be inversely correlated with colorectal cancer incidence. Per 90 grams of wholegrains per day, there was a significant 17% reduction in CRC risk (RR 0.83, 95% CI 0.78-0.89), concluding that consumption of wholegrains might be a protective factor against CRC.(50,54,55) Wholegrains are rich in various bioactive components, such as Vitamin E, copper, zinc and selenium that contain anti-carcinogenic properties, which might account for the inverse relationship between wholegrain consumption and CRC development.(50,56)

Dietary fibre is defined as undigested constituents of the plant cell wall and can be classified according to its source: cereal fibre, vegetable fibre and fruit fibre.(50) Various mechanisms have been proposed as protecting against CRC, such as the production of short-chain fatty acids with anti-proliferative effects and reduction in the intestinal transit time, resulting in reduced interaction of faecal mutagens with the colon mucosa. (54,57)

1.3.1.3 Alcoholic drinks

Many studies have investigated the relation of alcohol to CRC risk, and WCRF data observed a significantly higher risk for CRC following alcohol consumption above 30 grams daily (RR 1.15, 95% CI 1.06-1.26), which is equivalent to approximately two drinks per day. This relationship is significant with increased alcohol intake and the findings are significant when also stratified by sex. The mechanisms of how alcohol (such as ethanol) increases CRC risk have been well-established. Ethanol is metabolised into acetaldehyde, which is toxic to tissues, including colonocytes. Increased consumption of ethanol can also lead to the production of reactive oxygen species, which can lead to DNA damage.(50,58)

1.3.2 Body fatness and weight change

Body fatness can be measured by proxys, such as Body Mass Index (BMI). A BMI of 18-24.9 is considered to be normal, above that it increases to overweight and obese. The dose-response meta-analysis from 38 studies conducted by WCRF concluded that there is a significant 5% increased risk of CRC per 5 kg/m^2 (RR 1.05, 95% CI 1.03-1.07). The association appears to be stronger above 27 kg/m^2, classifying greater body fatness or obesity as a convincing cause of CRC. This can be attributed to the fact that body fat releases insulin, resulting in increased cell growth and inhibition of apoptosis both in human and experimental studies.(50,59)

Independent of body fatness, increase in weight is also associated with increased CRC risk. There is a 3% increased risk of CRC with a weight gain of 5kg and the association is stronger in men. Decrease in the weight through bariatric surgery has proven to decrease this risk by 27%.(60–62)

1.3.3 Physical activity

When comparing highest and lowest levels of physical activity, the former group showed a significant inverse association (RR 0.80, 95% CI 0.79-0.88),

decreasing the risk by 20%.(50) Increased physical activity results in reduction of insulin resistance and inflammation- both of which have been associated with tumorigenesis. Physical activity also stimulates digestion processes, reducing the transit time of faeces in the intestine; however, the evidence to support this in humans is limited.(63) Whether physical activity acts solely independently or whether the benefits are linked to loss of body fatness as well is not established.

1.3.4 Smoking

Smoking is a well-proven risk factor for many cancers, including CRC. A systematic review conducted 4 dose-response meta-analyses and showed significance associations with CRC risk, including daily cigarette consumption (RR = 1.38 for an increase of 40 cigarettes/day), duration (RR = 1.20 for an increase of 40 years of duration), pack-years (RR = 1.51 for an increase of 60 pack-years) and age of initiation (RR = 0.96 for a delay of 10 years in smoking initiation).(64) From all the CRC cases in the UK, 7% are attributed to tobacco smoking. (49)

1.4 Other CRC risk factors

1.4.1 Socioeconomic status

Socioeconomic status measures social, work and economic status, by measuring education, occupation and income respectively.(65) Deprivation can be measured differently, in Scotland the Scottish index of Multiple Deprivation, which encompasses education, employment, crime, housing and access to health services.(66) Low SES is considered a risk factor for CRC. The incidence of CRC is higher among males living in deprived areas in England. There is also higher mortality for males (30%) and females (15%) living in most deprived areas of England.(67) This trend can be attributed to the fact that behavioural risk factors, including poor diet, smoking and increased alcohol consumption are associated with deprivation. Even though awareness of the bowel screening programme is not affected, there is reduced screening uptake in these areas, with 66.6% uptake

in the least deprived compared to the 45.5% of the most deprived groups.(68,69) The decreased screening uptake can also partly explain the higher incidence.(69)

1.4.2 Diabetes

Diabetes has been implicated as a risk factor for many cancers.(70) CRC risk increases by around 30% in people with type II diabetes, according to meta-analyses.(70–73) It has also been proven that diabetes is positively associated with increased CRC mortality (RR 1.20, 95%, CI: 1.03-1.40).(71,72) These data are consistent when stratified according to sex and geographical location.(74) Furthermore, meta-analyses have shown that metformin users may have 11% lower risk of CRC than non-users.(75)

1.4.3 Inflammatory bowel disease

Inflammatory bowel disease (IBD) is the collective term used to describe ulcerative colitis and Crohn's disease. Due to the chronic inflammation of the bowel, a meta-analysis has shown that people with IBD have 70% higher risk of developing CRC compared to the general population.(76) The risk also increases with the extent and duration of IBD, as well as the location of the lesions.(76,77)

1.4.4 Aspirin

Aspirin has anti-inflammatory effects and some studies have shown aspirin to lower CRC risk by 32-49%.(78) In the meta-analyses of randomised control trials, aspirin has proven to be an effective chemopreventative agent, by reducing CRC risk by 17%.(79) The effects of aspirin may be more pronounced in people with higher BMI. This is because obesity is linked with increased inflammation thus, anti-inflammatory use might help in reducing this risk.(80)

1.5 Hereditary and Familial Colorectal cancer

1.5.1 Familial Colorectal Cancer

1.5.1.1 Background and aetiology

Approximately 30% of all colorectal cancer cases are thought to have some familial component. (81,82) However, highly penetrant inherited mutations and well characterised clinical presentations can only account for about 5% of those cases. Thus, the aetiologies of the remaining familial CRCs are currently not completely understood.(83) Familial CRC, also known as non-syndromic, can be defined as CRC clustering in families, which cannot be associated with well-known hereditary syndromes.

A number of different factors may contribute to the increased risk observed in these families. It has been suggested that higher familial risk to CRC can be caused by inheritance of mutations in single genes that are less penetrant than the genes causing the hereditary syndromes, but which are simultaneously more common. Another hypothesis for the familial clustering is the inheritance of multiple polymorphisms that result in an additive effect, also known as polygenic inheritance.(83,84) Shared environmental exposure may also contribute to the higher cancer aggregation in families, as there are a lot of factors previously mentioned that can modify CRC risk.

1.5.1.2 Genome-wide association Studies (GWAS)

There have been many genome-wide association studies that have identified common genetic risk loci for CRC.(85) A study that recruited 1,807 affected individuals and 5,511 controls found that variant rs6983267 on chromosome 8q24 was significantly associated with CRC (odds ratio = 1.22; $P = 4.4 \times 10^{-6}$).(86) Other studies have identified different single nucleotide polymorphisms (SNPs), also associated with increased risk, such as rs3802842 on 11q23 (OR = 1.1; $P = 5.8 \times 10^{-10}$), rs7014346 on 8q24 (OR = 1.19; $P = 8.6 \times 10^{-26}$) and rs4939827 on 18q21 (OR = 1.2; $P = 7.8 \times 10^{-28}$).(87,88) The results from the studies have been evaluated in a systematic review and showed that the reported

variants on the reported loci are considered to be highly credible but also identified 23 less credible variants at 22 loci.(84,89) **Figure 5** shows some identified SNPs and the risk of CRC for each.

Locus	Gene†	SNP	Per risk allele OR	Frequency risk allele
1p36.2	*WNT4; CDC42*	rs72647484	1.21	0.91
1q25.3	*LAMC1*	rs10911251	1.05	0.54
1q41	*DUSP10; CICP13*	rs6687758	1.09	0.2
2q32.3	*NABP1; MYO1B; SDPR*	rs11903757	1.06	0.36
3p14.1	*LRIG1*	rs812481	1.09	0.58
3p22.1	*RP11; CTNNB1*	rs35360328	1.14	0.16
3q26.2	*MYNN; TERC*	rs10936599	1.08	0.75
4q26	*NDST3*	rs3987	1.36	0.44
4q32.2	*FSTL5*	rs35509282	1.53	0.09
5q31.1	*PITX1; H2AFY*	rs647161	1.11	0.67
6p21.31	*CDKN1A*	rs1321311	1.1	0.23
8q23.3	*EIF3H*	rs16892766	1.25	0.07
8q24.21	*CCAT2; MYC*	rs6983267	1.21	0.52
9q24	*TPD52L3; UHRF2*	rs719725	1.19	0.37
10p13	*CUBN*	rs10904849	1.14	0.68
10p14	*GATA3*	rs10795668	1.12	0.67
10q22.3	*ZMIZ1; AS1*	rs704017	1.06	0.57
10q24.2	*SLC25A28; ENTPD7; COX15; CUTC; ABCC2*	rs11190164	1.09	0.29
10q25	*VTI1A*	rs12241008	1.13	0.09
11q12.2	*FADS1; FEN1*	11qhap‡	1.4	0.57
11q13.4	*POLD3*	rs3824999	1.08	0.5
11q23.1	*COLCA2*	rs3802842	1.11	0.29
12p13.32	*CCND2*	rs3217810	1.2	0.16
12p13.32	*CCND2*	rs3217901	1.1	0.41
12p13.32	*CCND2*	rs10774214	1.09	0.38
12q13.13	*DIP2B; ATF1*	rs11169552	1.09	0.72
12q13.13	*LARP4; DIP2B*	rs7136702	1.06	0.35
12q24.12	*SH2B3*	rs3184504	1.09	0.53
12q24.21	*TBX3*	rs59336	1.09	0.48
12q24.22	*NOS1*	rs73208120	1.16	0.11
14q22.2	*BMP4*	rs1957636	1.08	0.4
14q22.2	*BMP4*	rs4444235	1.11	0.46
15q13.3	*SCG5; GREM1*	rs11632715	1.12	0.47
15q13.3	*SCG5; GREM1*	rs16969681	1.18	0.09
16q22.1	*CDH1*	rs9929218	1.1	0.71
16q24.1	*FOXL1*	rs16941835	1.15	0.21
17q21	*STAT3*	rs744166	1.27	0.55
18q21.1	*SMAD7*	rs4939827	1.18	0.52
19q13.11	*RHPN2*	rs10411210	1.15	0.9
19q13.2	*TMEM91; TGFB1*	19qhap‡	1.16	0.49
20p12.3	*FERMT1; BMP2*	rs2423279,	1.14	0.3
20p12.3	*FERMT1; BMP2*	rs4813802	1.09	0.36

Figure 5 *Summary description of meta-analysis of identified SNPs*. Figure taken from source 85.

1.5.1.3 Risk classification

Different studies have investigated the risk of colorectal cancer and adenomatous polyp development in familial clustering. These studies showed that there is a higher risk of polyp and cancer development compared to the general population, but lesser risk compared to the Lynch syndrome group (*See* **Section 1.5.2**). The personal risk is greater with increasing number of affected relatives and decreasing age of onset of those relatives.(90–92) A systematic review and meta-analysis concluded that RR of CRC doubled for individuals with at least 1 FDR with CRC and almost tripled with two or more FDR's. The RR also tripled when the relative was less than 50 years old at diagnosis, in contrast with people with a FDR more than 50 at time of CRC diagnosis, with a cumulative absolute risk estimate at age 85 of less than 5%.(92–94) Combining data from these studies, the BSG have classified patients with different family history of colorectal cancer into risk groups, presented in **Table 5.**(37) In the low-moderate groups, Lynch syndrome should not be excluded just on the basis of the Family History criteria, so immunohistochemistry (IHC) or microsatellite instability (MSI) testing should be carried out from pathology tumour material if available.

Table 5 *Risk stratification according to family history of colorectal cancer.* Information taken from source 37.

Risk group	*Family History Criteria*
Moderate-Low risk	• Colorectal cancer in 1 FDR <50 years • Colorectal cancer I n 2 FDR ≥ 60 years old
Moderate- High	• Colorectal cancer in 3 FDR in first degree kinship, none <50 years • Colorectal cancer in 2 FDR in first degree kinship, mean age <60 years
Low risk	• Other Family History of Colorectal Cancer

First Degree Relative: FDR; first degree kinship: Affected relatives who are first-degree relatives of each other AND at least one is a first degree relative of the consultant. No affected relative <50 years old (otherwise high-risk criteria would apply). Combinations of 3 affected relatives in a first-degree kinship include: parent and aunt/uncle and/or grandparent; OR 2 siblings/1 parent; OR 2 siblings/1offspring. Combinations of 2 affected relatives in a first-degree kinship include a parent and grandparent, or >2 siblings, or >2 children, or child + sibling. Where both parents are affected, these count as being within the first-degree kinship.(37)

1.5.2 Hereditary non-polyposis Colorectal cancer (HNPCC)- Lynch Syndrome

1.5.2.1 Background and inheritance patterns

HNPCC accounts for 3% of inherited CRC cases and is the most common inherited condition for CRC.(95) In the UK, an estimate of 1200 cases per annum of CRC are attributed to this condition.(96) The condition was described in the early 20th century, where families were identified with increased cases of CRC. These cases shared a number of clinical characteristics including early age, improved clinical outcome and involvement of the right colon. Histologically, the cancers more frequently are poorly differentiated and mucinous. They have a large number of tumour-infiltrating lymphocytes and a high level of microsatellite instability.(97) However, the term can be misleading due to the polyp formation extra-colonic cancers.(95,97,98) Malignancies involving the ovaries, endometrium, stomach, hepatobiliary epithelium, small bowel, uroepithelial epithelium and brain are associated with HNPCC.(99) After colorectal cancer risk, the most common malignancy in HNPCC women is in the endometrium followed by the ovary with a lifetime risk of 27-71% and 3-14% respectively followed by the rest of the cancers (*See* **Figure 6**).(100–102) Henry Lynch elaborated the condition by marking the mode of inheritance and the tendency for the right colon; the condition was replaced with the term Lynch syndrome in cases linked with a germline pathogenic variant.(95,97,98) Lynch syndrome is inherited in an autosomal dominant manner, meaning that there is 50% chance of any child to inherit the disease if they have an affected parent.(95)

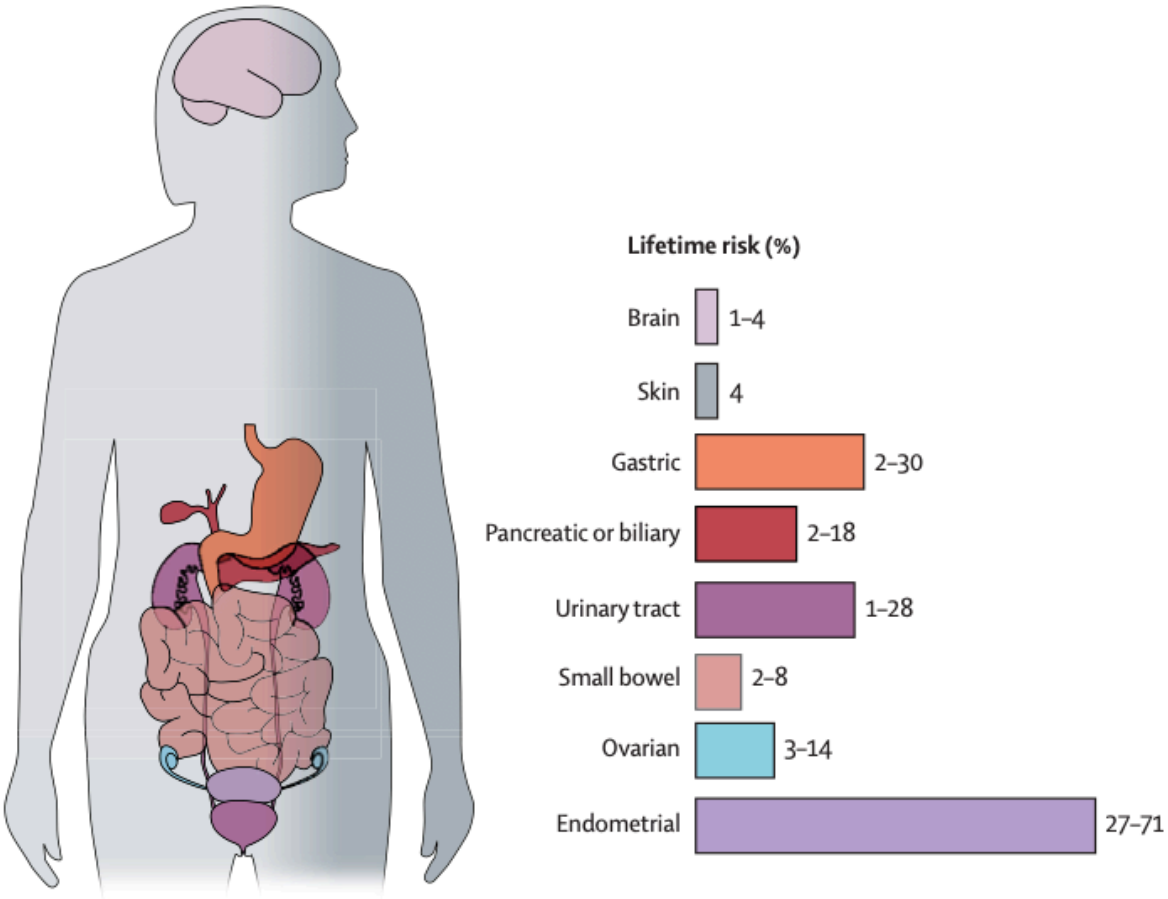


Figure 6 *Lifetime risk of extra-colonic cancers associated with Lynch syndrome.* Figure taken from source 102.

1.5.2.2 Associated genes

The aetiology of HNPCC has been associated with changes in several genes (*See* **Table 6**).(103) These genes include the mismatch repair (MMR) genes, such as MLH1, MSH2, MSH6 and PMS2.(97) These genes encode proteins that are involved in DNA repair. Mutations in such genes results in a defective MMR mechanism, thus allowing errors in the DNA to be left unrepaired and increase substantially during the cell cycle. Accumulating errors result in abnormal function of the cells, increasing the risk for colon and other tumours.(104) MMR defects are associated with the molecular phenotype of Microsatellite Instability (MSI) in tumour DNA, defined as alternate sized repetitive DNA sequences that are not present in the corresponding germline DNA.(105) In Tayside, all colonic tumour pathology specimens are tested for MSI, as it is a hallmark of MMR defects. Nevertheless, the presence of an isolated MSI in tumour together with MLH mutation should be interpreted with caution, as promoter hypermethylation which results in MLH1 loss occurs frequently via epigenetic silencing, due to somatic mutations.(106)

Another gene associated with Lynch syndrome is the EPCAM gene.(107) Although EPCAM gene is not directly involved with DNA repair, it lies next to the MSH2 gene on chromosome 2. Mutations in the EPCAM gene can cause hypermethylation of the promoter regions, resulting in decreased activity or complete deactivation of the MSH2 gene.(104)

1.5.2.3 Increased risk

There is an 80% lifetime risk of developing CRC with Lynch syndrome, with the average age of onset being less than 45 years; this varies according to specific mutations on different genes. The risk also varies between the two genders, with males having a greater risk of developing CRC than females.(37,108,109) In order to identify patients that might have inherited one of those genes, the Amsterdam I criteria were developed (*See* **Section 1.6.1**).(95) However, more than half of the cases were missed, leading to the revised Bethesda guidelines. When patients meet these criteria, they go on for further evaluation and genetic

testing. Another approach in identifying Lynch syndrome is tumour testing when Bethesda guidelines are identified.(106) This has proved to be cost effective and involved with testing for MSI and/or immunohistochemical analysis of CRC tumours.(110) Even though this analysis is very sensitive as approximately 90% of these cancers will have an MSI, the specificity is lower as 15% of sporadic CRCs are also MSI-H. IHC is conducted by utilising four antibodies specifically related to the MMR genes to evaluate tumours for MMR deficiency. Tumour testing can also be performed on endometrial cancers in order to effectively identify Lynch syndrome; tumour testing on other related tumours has not been evaluated.(83) Other models such as PREMM, MMRpro, and MMRpredict have also been developed and are discussed in **Section 1.7**.

Table 6 *Summary table of HNPCC-associated genes*. The table shows the genes associated with HNPCC, along with the percentage of cases attributed to each gene. It also shows the risk of developing HNPCC cancers when inheriting mutations from each. Information extracted from sources 104 and 109.

Gene	Chromosomal location	Lynch cases attributed to the gene (%)	**Lifetime cancer risks**						
			Colon	Endometrial	Stomach	Ovarian	Small bowel	Hepatobiliary tract	Urinary tract
MLH1	3p22.2	50%	22-74%(males) 22-53%(females)	14-54%	0.2-13%	4-20%	4-12%	0.4-4%	0.2-25%
MSH2	2p21-p16.3	40%	22-74%(males) 22-53%(females)	40-60%	0.2-13%	4-20%	4-12%	0.4-4%	0.2-25%
MSH6	2p16.3	7-10%	22% (males) 10% (females)	16-26%	6-22%	22%	6-22%	6-22%	6-22%
PMS2	7p22.1	<5%	20% (males) 15% (females)	15%	6%	6%	6%	6%	6%

1.5.3 Familial adenomatous polyposis

1.5.3.1 Background and inheritance pattern

Familial adenomatous polyposis (FAP) is the second most common inherited disorder that increases CRC risk, with prevalence from 1 in 7000 to 1 in 22000 people, occurring most commonly in the western countries.(95,111,112) The condition was increasingly reported in the 1900's, describing multiple polyp formation which occurred in autosomal dominant pattern; the adenoma to carcinoma progression was later on confirmed. FAP is the most clearly defined inherited colon cancer; it is characterised by the development of multiple adenomatous polyps (>100) in the colon and rectum, after the first decade of life. Furthermore, polyps might appear in the upper GI tract such as in the stomach and duodenum. The condition is associated with extraintestinal features as well, such as osteomas, epidermoid cysts, desmoid tumour formation, supernumerary teeth and Congenital hypertrophy of retinal pigment epithelium (CHRPE). Other tumours may also arise, such as hepatoblastoma and brain tumours, most commonly medulloblastoma.(95) The risk of developing adenomas is around 90%, with a median age of 16 years old.(95,113) This risk rises steeply with age, with manifestation of adenomas at 10 years being 15%, rising to 75% by the age of 20 and to 90% by the age of 30.(113,114) Without surveillance and intervention, APC mutation carriers will develop CRC by the fourth decade of their life.(95,113) Surveillance is vital for the early detection of colonic polyps in order to prevent CRC with colectomy.(37,95,115)

1.5.3.2 Associated genes

Mutations on the Adenomatous Polyposis Coli (APC) gene, located on chromosome 5q21 are responsible for FAP.(116) APC is a tumour suppressor gene and codes for APC protein that is important for cell adhesion and signal transduction. An example is the role of APC in the signalling of beta-catenin break-down when it is not needed.(95,111) Beta-catenin helps control the expression of genes that promote proliferation and differentiation; thus, mutations in the APC genes would result in uncontrolled proliferation and differentiation due to defective APC signalling. More than 300 pathogenic mutations have been reported in APC primarily in the first half of

the gene (codons 169 to 1393).(95,117,118) The identified mutations are insertions, deletions and nonsense variants that create premature stop or frameshift mutations that lead to truncated protein. The variant that dominates FAP pathogenicity is the deletion of AAAG in codon 1309, which occurs in 10% of the FAP patients.(119) Pathogenic variants of the APC gene result in approximately 90% chance of getting the disease; both in men and women.(95)

1.5.3.3 Attenuated Adenomatous Polyposis

AFAP is a variation of FAP which predisposes to an increased risk of CRC, however, with a lower number of adenomas (average 30) and with predominantly right sided polyps.(120) In contrast to FAP, adenomas begin to form in the late twenties and CRC risk is higher at a later stage, at an average age of 56.(118,121,122) The extra-colonic manifestations are similar to classic FAP with the absence of CHRPE lesions. A subset of APC pathogenic variants is associated of AFAP. Pathogenic variants at the 5' end of the APC gene and exon 4 variants usually present with 2-500 polyps.(118,123) Moreover, exon 9 pathogenic alterations result in the formation of 150 adenomatous polyps or less and no upper GI manifestations. Region 3' pathogenic variants lead to approximately 50 adenomas.(123)

1.5.4 MUTYH-Associated polyposis

MUTYH-Associated polyposis (MAP) is an autosomal recessive inherited syndrome, also characterised by the presence of adenomatous polyps in the colon, leading to an increased risk of CRC. The risk of CRC for MAP patients is lower than FAP, but still ranges from 35-75% throughout life.(95,124) Clinically, MAP resembles attenuated FAP; however, patients with this inherited syndrome develop fewer adenomas than the ones with APC pathogenic mutations.(125–127) Colonic polyposis typically occurs by the age of 40, although it can emerge at earlier ages.(128) Furthermore, colonic cancers tend to be right sided and there seems to be better prognosis than sporadic CRC.(129) MAP is mostly associated with hyperplastic polyps (47%) and serrated adenomas.(130) Extra-intestinal features include different types of cancer, such as gastric, small intestinal, endometrial, breast, liver, ovarian, bladder and thyroid. Skin cancers have also been reported, including melanoma, squamous epithelial and basal

cell carcinomas. Non-cancerous features include lipomas, osteomas and desmoid tumours.(131,132) The incidence of the extracolonic manifestations is lower in patients with MAP relative to patients with FAP or Lynch.(129)

The gene associated with MAP is MUTYH gene, also referred as the MYH gene. This gene was firstly identified in 2002 and its cytogenic location is 1p34.1.(133) The MUTYH gene codes for the enzyme MYH glycosylase which is involved in DNA repair. MYH glycolysis has a role in base excision repair, preventing G:C to T:A transversions caused by oxidative stress. The inheritance pattern of MAP is autosomal recessive, meaning that both copies of this gene need to be mutated. This results in non-functional or low functioning MYH glycolysis, leading to impaired base excision repair mechanism. Subsequently, this leads to building up of mutations, leading to cell overgrowth and tumour mutations. The most common mutations associated with MUTYH gene are at position 179 (Tyr179Cys) and 396 (Gly396Asp).(133) Founder pathogenic variants, meaning mutations that occur in specific ethnicities, are assumed for MUTYH, such as Y179C and G396D which account for 70% of biallelic pathogenic variants of norther European MAP patients, P405L in Netherlands and E490X in India.(134–136) A study with 7225 individuals with colorectal adenoma reported a prevalence of 4% (95% CI, 3%-5%) and 7% (95%, CI 6%-8%) among patients with 10-19 and 100 to 999 adenomas respectively for biallelic MUTYH pathogenic variants. This implied that mutations in this gene are able to cause disease in both homozygous or compound heterozygous forms.(95)

1.5.5 Rare Colon Cancer Syndromes

1.5.5.1 Juvenile polyposis syndrome (JPS)

JPS is a disease, usually presenting in childhood or early adulthood, characterised by hamartomatous polyposis throughout the GI tract with predominating colorectal polyps.(137) It has an autosomal dominant inheritance pattern, with 75% of the cases being inherited from one affected parent and the remaining cases resulting from *de novo* mutations. The prevalence of JPS is one in 100 000 individuals.(95) JPS is diagnosed when someone meets one or more of the following criteria: more than five juvenile polyps of the colon or rectum, juvenile polyps in other parts of the GI tract, any

number of juvenile polyps and a positive family history of JPS.(138) Juvenile polyp is a specific type of hamartomatous polyp, based on histological appearance. Even though these polyps are benign, there is a 10 to 50% increased risk of people with JPS to develop cancer of the GI tract at some point in their lives.(139) Presentation usually comprises non-specific GI symptoms such as diarrhoea or GI tract haemorrhage. Patients may also have signs and symptoms of Hereditary haemorrhagic telangiectasia (HHT) such as arteriovenous malformations, digital clubbing and osteoarthropathy, suggesting an overlap of these two syndromes. Additionally, 15% of individuals with JPS present with other abnormalities such as cleft palate, polydactyly, intestinal malrotation and abnormalities in heart, brain, genitalia or urinary tract. JPS can be subdivided into three different types. The first one occurs in infancy, is the most severe form and is associated with the poorest outcomes and increased morbidity and mortality.(140) Symptoms include protein-losing enteropathy which results in severe diarrhoea, failure to thrive and cachexia.(139,141) The second type is Generalised JPS which is characterised by polyps developing throughout the GI tract, and the third type is called juvenile polyposis coli and affects solely the colon.(139)

Mutations in BMPR1A and SMAD4 genes are responsible for JPS.(142,143) The BMPR1A is located at 10q23.2 and codes for the bone morphogenetic protein receptor 1A.(143) This protein binds to a ligand, activating a protein complex called SMAD proteins which are then transported to the cell's nucleus to regulate proliferation and the activity of other genes. Mutations in the BMPR1A gene produces an abnormally short, non-functional protein, unable to ligand and activate the SMAD complex, resulting in unregulated cell growth that can lead to polyp formation. More than 60 pathogenic variants, including nonsense, frameshift, missense and splice-site variants of BMPR1A have been identified and account for 25-40% of JPS cases.(139,143,144) SMAD4 is located on the long arm of chromosome 18, at position 21.2 and is implicated in 15-60% of JPS cases.(137) SMAD4 acts both as a transcription factor and a tumour suppressor and it is part of the transforming growth factor beta (TGB-β) pathway which regulates cell proliferation. Apart from JPS, SMAD4 mutations have been implicated in some cancers and conditions such as HHT, hence the overlap with JPS.(142,145) Until today, 78 pathogenic variants that lead to JPS have been identified between exons 6 and 11.(95)

1.5.5.2 Peutz-Jeghers syndrome (PJS)

PJS is an early onset autosomal dominant disorder, characterised by the development of both benign hamartomatous and adenomatous polyps in the GI tract and increased risk of multiple cancers.(146) Signs of PJS are small, dark coloured spots on the lips and inside the mouth, near the eyes and nostrils and around the anus. There is also increased development of polyps that can cause recurrent bowel obstructions, chronic bleeding and abdominal pain. PJS predisposes to an increased risk of GI tract cancers, as well as pancreas, cervix, ovary and breast.(95,146) The estimated prevalence is 1 in 25000 to 30 000 individuals with a cumulative risk of around 40% by the age of 65 for CRC.(147)

Mutations in the STK11 gene result in most cases of inherited PJS syndrome. This gene is located at 19p13.3 and acts as a tumour suppressor gene, hindering uncontrolled growth and proliferation.(148) Studies have demonstrated the development of hamartomas in heterozygous STK11 knock-out mice without the inactivation of the wild type allele. This suggests that initial tumour development in PJS can occur with haploinsufficiency (STK11 +/-) due to loss of heterogeneity (LOH).(149) More than 340 pathogenic variants have been associated with PJS, producing a short non-functional serine/threonine kinase 11 enzyme which impairs its function.(148) The mutations are mostly localised to regions that code for the kinase domain of the protein and include a variety of nonsense, frameshift, missense, splice-site variants and large deletions.(150)

Table 7 *Summary of inherited conditions that increase risk of CRC development.*

Syndrome	Inheritance	Genes	Functions	Phenotype	Risk of CRC	Frequency in CRC
Non-polypotic syndromes						
HNPCC	AD	MLH1 MSH2 MSH6 PMS2	DNA mismatch repair	Early onset CRC Increased risk of extracolonic cancers such as endometrial, ovarian, gastric	50-80%	2-4%
Polypotic, adenomatous						
FAP	AD	APC	Cell adhesion and signal transduction	100-1000 polyps Duodenal and bowel adenomas Upper gastrointestinal cancer risk	100%	1%
AFAP	AD	APC	Cell adhesion and signal transduction	Milder phenotype, 0-100 polyps Duodenal and bowel adenomas Upper gastrointestinal cancer risk	75%	1%
MAP	AR	MUTYH	DNA base excision repair	<100 polyps Serrated adenomas and hyperplastic polyps Colon cancer, rarely gastric cancer	35-75%	<1%
Polypotic, hamartomatous						
JPS	AD	BMPR1A SMAD4	Regulation of cell proliferation through TGB-β signalling	Multiple polyps in colon and throughout gastrointestinal tract	10-50%	<1%
PJS	AD	STK11	Tumour suppressor, proliferation control	Small polyps in bowel, small intestine and stomach Oesophageal, gastric, small intestine, colon and pancreatic cancer	40-70%	<1%

Autosomal Dominant (AD); Autosomal recessive (AR), Colorectal Cancer (CRC); Hereditary Non-Polyposis Colorectal Cancer (HNPCC); Familial adenomatous polyposis (FAP); Attenuated FAP (AFAP); MUTYH associated polyposis (MAP); Juvenile polyposis syndrome (JPS); Peutz-Jeghers syndrome (PJS)

1.6 Risk Assessment models

A variety of models have been created in order to predict risk of carrying the pathogenic mutations related to Lynch syndrome. Even though all the models have the same purpose, they differ in terms of the variables used to predict the risk as well as the way that they were developed. In addition, each one has its own advantages and disadvantages, in terms of sensitivity, specificity, cost-effectiveness and accessibility. However, the fact that different populations were used to validate these models might impact on their accuracy. Thus, when deciding which specific model to use, both the patient population that is being evaluated as well as the clinical setting must be taken into consideration.

1.6.1 Amsterdam criteria

In 1990, the International Collaborative Group meeting in Amsterdam developed a set of criteria in order to identify families likely to have Lynch syndrome, known as the Amsterdam criteria.(95) The criteria included having at least three relatives with CRC and meeting all of the following : one affected individual is a first degree relative of the other two; at least two successive generations are affected; at least one CRC is diagnosed before the age of 50 years; FAP should be excluded and tumours verified by pathological examination. However, these criteria were not sensitive enough as more than 50% of families with Lynch syndrome failed to meet these criteria.(151) Thus, Amsterdam criteria II were developed in 1999.(152) These criteria had as core principles not to deviate largely from the initial criteria and focus on clinical aspects as genetic testing might not be accessible to all families. The criteria remained the same with the addition of a new criterion: at least three relatives should have an HNPCC-related cancer (CRC, endometrium, small bowel, ureter, renal or pelvis).(152) When the causative mutations were identified, these criteria were proven to be specific but not sensitive predictors of MMR gene carriers.(153) On further evaluation, families have been identified that do not fulfil the criteria but were diagnosed with HNPCC.(154) Nevertheless, these criteria are still categorised as level IV evidence, coming from expert committee reports.(37) Thus, Amsterdam II criteria can still be used as a clinical tool to identify families with risk to carry DNA MMR

gene mutations.(155) Since the absolute risk of CRC development with Lynch can rise up to 56% depending on the mutated gene, any presentation of CRC before the age of 50 should be investigated for a possible Lynch diagnosis.(37,95) It should also be noted that in around 20% of the families that meet the Amsterdam criteria and exhibit MSI or loss of DNA MMR gene, the germ-line mutation cannot be detected using current methodologies.(95)

For very small families, modified Amsterdam criteria are applied, thus a person can be considered to be at risk of HNPCC with only two CRCs in FDR, in which one of the cancers was diagnosed before the age of 55. If two FDRs are affected by CRC, a third relative with an early onset HNPCC related cancer such as in the endometrium is sufficient for meeting the criteria. In patients with a very early onset of CRC diagnosis (before 40), with no FH, the patient is still considered to be at risk of having HNPCC.(156)

1.6.2 Bethesda guidelines

In 1997, the National Cancer Institute met in Bethesda in order to discuss identification of people with HNPCC.(106) This led to the development of a set of guidelines, aiming to categorise people with HNPCC that should be tested for MSI, as MSI accounts for approximately 15% of CRC.(95) The original guidelines included a panel of genes that were tested in order to identify MSI: BAT25, BAT26, D2S123, D5S246 and D17S250. However, these guidelines came with a set of limitations, so in 2002, another workshop was held in order to tackle issues that were identified and consider revision and improvement (*See* **Table 8** for the revised guidelines).(106) People identified as high-risk are then recommended to have molecular evaluation, either through MSI or IHC analysis of tumour followed by germline testing or directly through germline testing of mutations. At-risk relatives should be given the option of genetic counselling and testing. If a mutation is not identified but clinical suspicion of HNPCC is high, then patients and their at-risk relatives should be counselled and commence surveillance.(106) Members of the workshop included Dr Henry Lynch and Dr Albert Warthin, who were the first who suspected and later on discovered this syndrome.(97) In order

to identify the sensitivity and specificity of the modified Bethesda guidelines, a study interviewed 127 CRC patients who were considered to be high-risk based on these guidelines. The investigators reviewed their medical records as well as the performed MSI analysis of tumours. From the 127 patients, 42% of those were found to have MSI-H tumours. Interestingly, 36 patients were tested for mutations with 61% of them testing positive. Significant predictors of MSI were early age of CRC diagnosis, number of CRC per family and presence of other HNPCC cancers in the family. Presence of multiple cancers in a single-family member was a specific predictor of MSI status, regardless of age.(157) Further studies revealed that these guidelines are 96% sensitive for identifying MSI-H tumours in high risk populations, however, their specificity is relatively low (27%).(158) Nevertheless, the aim of the guidelines is not to identify MSI tumours from patients in the general population, but HNPCC patients. This means including MSH2 and MLH1 mutation carriers. When investigating specificity and sensitivity of these criteria in MSH2 and MLH1 mutation carriers, the Bethesda guidelines were found to be the most sensitive but the least specific compared to the existing criteria (Amsterdam I, II and modified Amsterdam criteria).(159) These guidelines have also proven to be highly cost-effective when identifying HNPCC.(95,106)

Table 8 *Revised Bethesda Guidelines*. Information extracted from source 106.

The Revised Bethesda Guidelines for testing colorectal cancer tumours for MSI
1. CRC diagnosed in a patient who is less than 50 years of age 2. Presence of synchronous, metachronous colorectal or other HNPCC-associated tumours*, regardless of age 3. Colorectal cancer with the MSI-H histology diagnosed in a patient who is less than 60 years of age 4. Colorectal cancer diagnosed in one or more FDR with HNPCC- related tumour with one of the cancers being diagnosed under the age of 50. 5. Colorectal cancer diagnosed in two or more FDR or SDR with HNPCC-related tumours, regardless of age.
*CRC, endometrial, stomach, ovarian, pancreas, ureter and renal pelvis, biliary tract and glioblastoma of the brain, sebaceous gland adenomas and keratoacanthomas and carcinoma of the small bowel.

1.6.3 PRediction model for gene mutations (PREMM)

The $PREMM_1$ model is a clinical prediction algorithm that was introduced in 2006 and predicts the probability of carrying MLH1 and MSH2 gene mutations. Later on, $PREMM_{1,2}$ calculator provided the cumulative probability of carrying either of these two mutations. The PREMM1,2,6 model was introduced in 2011, replacing the previous two models as it incorporated their algorithms along with the cumulative probability of identifying MSH6.(160) The most recent development of PREMM5 in 2017 includes PMS2 or EPCAM gene mutations too.(161) In order to evaluate risk, the model requires personal or family history of CRC, endometrial or other Lynch syndrome-associated cancers. It also includes specific types of cancer and ages at diagnosis of both first- and second-degree relatives from the affected side of the family.(161,162) Advantages of this model is the fact that it is easily accessible via the web and it is simple in its use. It has also been validated with a cohort of 1058 patients with CRC and includes a broad spectrum of extra-colonic cancers.(162) Even though the risk prediction model includes first- and second-degree relatives, it does not take into account family size and thus may overestimate the risk in some cases.(95,163)

1.6.4 MMRpro

MMRpro is another statistical model which was developed to assess the probability of carrying MLH1, MSH2 and MSH6, based on family history of CRC and endometrial cancer.(164) A major disadvantage is the fact that it does not include PMS2 and EPCAM. Moreover, its use is restricted to specialists due to its limited access. It has been also deemed as more time consuming relative to other models as it requires data from an individual's entire pedigree.(163–165)

1.6.5 MMRpredict

MMRpredict also uses a statistical model to identify patients with mismatch repair gene mutations. The criteria used are age, gender, location of tumour, personal history of CRC and other cancers, family history of CRC and other cancers.(166) However, this model has been found to be less accurate as the model was

developed only by using CRC patients of less than 55 years of age and did not include extracolonic malignancy.(164)

1.6.6 Polygenic risk and personalised screening

A number of studies have attempted to personalise the screening strategy of CRC by using FH data as well as polygenic risk scores (PRS).(167–169) The latter refers to known SNPs, identified in GWAS studies that have proven to increase the risk of CRC development.(170) There is an increasing number of SNPs associated with CRC, and although the risk associated with each SNP is small, combination of multiple SNPs in genetic risk scores may be clinically relevant and allow more targeted CRC prevention and early detection.(167) However, the cost-effectiveness of this type of screening is not yet known and further comparative studies between PRS and current guidelines are required before clinical utilisation.

1.7 Guidelines and Surveillance Recommendations

1.7.1 Moderate risk family history

Since there is an increased risk of polyp and cancer development in people with non-syndromic familial CRC, increased surveillance is recommended to minimise this risk through colonoscopic screening and adequate polypectomy. Studies comparing familial colorectal cancer in families with and without Lynch syndrome has shown that the risk of polyps and cancer is lower in people without Lynch syndrome.(91,171–174) They have also identified that people under the age of 45-50 years may not require any surveillance if there is not an extensive family history of CRC and three yearly colonoscopies in people with an extended family history.(171,172,174)

Based on these studies, different countries have given recommendations on surveillance guidelines.(37,175) Currently, there are no guidelines issued by the National Institute of Health and Care excellence (NICE), but the BSG have published their recommendations based on these studies and expert opinion,

which are also utilised by Scottish Intercollegiate Guidelines Network (SIGN).(176) For low-moderate risk patients (as classified by the BSG), the recommendation is once only colonoscopy at 55 years of age with no follow-up if this is normal. For people in the high-moderate group, BSG recommends colonoscopic surveillance to start at age 50, and repeat every five years until the age of 75.(37) Standard surveillance guidelines will be used if the colonoscopy is abnormal, for example adenomatous polyp detection (*See* **Section 1.2**). These guidelines fail to include patients with extensive family history of high-risk adenomas without a clinical and family history of polyposis conditions. In contrast, the US multi-society task force on colorectal cancer includes guidelines on people that have relatives with advanced adenoma, depending on number of polyps and age of relatives. In general, these guidelines recommend screening at an earlier age, for example if a person has a FH of CRC or a FDR with advanced adenoma age <60 or 2 FDR with CRC at any age, the recommended surveillance is colonoscopy every 5 years beginning 10 years before the age of diagnosis of the youngest relative or age 40, based on whichever is earlier. They also include recommendations for people with a single FDR with CRC diagnosis ≥60, again starting screening at age 40.(175)

1.7.2 Lynch syndrome

Compared to moderate risk groups, people with Lynch syndrome have a higher risk of cancer and adenomatous polyp development. These has been shown in many studies, with a 1 in 5 and a 1 in 13 risk in males and females respectively to develop CRC if they fulfil Lynch syndrome criteria. Most studies recommend that the ideal surveillance time would be every one or two years, depending on the study.(91,172,174,177) For people at risk of HNPCC that fulfil the modified Amsterdam criteria or they are an untested FDR of a proven mutation carrier or they are an MMR gene carrier, colonoscopy is advised from age 25, every 18-24 months, as well as upper gastrointestinal endoscopy of the oesophagus, stomach and duodenum (OGD) from the age of 50, every two years. People with 1 FDR with MSI-H CRC and IHC loss of MSH2, MSH6 or PMS2 expression, colonoscopic surveillance is also advised to start at 25 years of age and be repeated every two years with two-yearly OGD starting at 50 years old.(37) As

mentioned, MSI-H CRC with IHC loss of MLH1 is excluded in elderly patients with right sided tumours as it is usually a somatic epigenetic event.(37) Surveillance would also continue in the event of partial colectomy, usually with two yearly sigmoidoscopies.

1.7.3 FAP

Due to the high risk of carcinoma development in an early age, FAP are strongly advised to have a procto-colectomy and pouch or colectomy before the age of 30. In patients at risk of FAP but without an identified mutation, colonoscopy or alternating colonoscopy with flexible sigmoidoscopy should be commenced at puberty with annual screening until the age of 30 and thereafter 3-5 years until 60 years. Procto-colectomy or colectomy is favourable if the patient is clinically positive. If the person fulfils the FAP criteria or is a proven APC mutation carrier opting for a deferred prophylactic surgery, colonoscopy or alternating colonoscopy with flexible sigmoidoscopy as well as OGD is recommended to start at diagnosis or puberty twice a year.(37) Puberty is loosely defined to allow a flexible approach, depending on the level of maturity of the individual. The guidelines also do not differentiate between FAP and AFAP, in which the latter has a much later age of colorectal cancer incidence risk, as well as fewer number of polyps developing. Patients after colectomy and or ileorectal anastomosis are still advised to be screening with annual flexible or rigid rectoscopy or pouch endoscopy (depending on the procedure) as well as OGD every three years. This is because studies have shown risk of rectal cancer to be around 20% after total colectomy.(178,179) Screening recommendations for other extra-colonic cancers are beyond the scope of this study and can be accessed through different guidelines.(180)

1.7.4 Other High-risk categories

People with MAP are recommended to have colonoscopic screening starting from age 25 biennially and OGD from age 30, 3-5 yearly.(37) Similarly, to FAP, mutation carriers should be individually counselled for prophylactic surgery options. The European society of medical oncology (ESMO) also has similar

recommendations.(180) For PJS, the colonoscopic advice is the same as MAP, but also there is an additional screening of small bowel video capsule endoscopy (VSE) or MRI every 2 to 4 years. Screening for stomach cancer is also different, with OGD starting at age 25 and repeating every two years.(37) The National Comprehensive cancer network (NCCN) has similar guidance for PJS, whereas ESMO guidelines recommend the first colonoscopy at age 8 and the second at age 18 if the findings of the initial screening are normal.(37,180,181)

1.8 Basis of this research

The above findings clearly demonstrate the demand for the development of a revised risk model which is both sensitive and specific in identifying familial CRC in the clinical setting. According to stratified risk, this model should provide more information regarding the management of these patients and the screening recommendations.

The BSG guidelines recommend an audit regarding people attending the service who are concerned about their risk of CRC with outcomes including extent of family history, assignment of risk, and surveillance prevalence of cancer, adenomas as well as morbidity and mortality. To our knowledge, no one has attempted to validate the effectiveness of BSG guidelines, which are also recommended by the SIGN guidelines, in a patient cohort who have attended clinical genetics regarding their risk. In Tayside, it is feasible to analyse a large retrospective cohort of people who have attended clinical genetics and determine their BSG risk category and outcomes over a period of time. This pilot study aims to assess the effectiveness of risk stratification of the BSG guidelines and evaluate the recommended colonoscopic surveillance in the Tayside cohort, as well as demonstrate the methodology which could be used to conduct similar studies in other centres.

Chapter 2: Aims and objectives

2.1 Aims

The aims of this study are:

- To examine how effective the BSG guidelines for familial colorectal cancer are at identifying people at increased risk of colorectal cancer
- To identify which elements of family history, appear to predict an increased risk of colorectal cancer
- To investigate the relative and absolute risks of developing colorectal cancer depending on the risk category in the Tayside cohort
- To explore whether people with increased risk of CRC also have an increased relative and absolute risks of having adenomatous polyps being detected
- To evaluate whether guidelines identify people who will benefit from increased screening

2.2 Objectives

In order to achieve the above aims, the following objectives will be carried out:

- A cohort of patients who have attended the clinical genetics regarding risk of colorectal cancer will be identified and assigned to a risk category depending on BSG criteria. Those who subsequently develop colorectal cancer or have adenomatous polyps detected will be identified.
- Based on the collected information, risk of colorectal cancer and polyp risk for people in each category will be calculated.
- Statistical analysis will be performed to assess if any particular element of FH is significantly related to colorectal cancer risk.

Chapter 3: Methods

3.1 Approvals and data collection

3.1.1 Approval

NHS Tayside Caldicott approval (*See* **Appendix 1**) was obtained, to collect data concerning the cohort of patients that attended the clinical genetics department with a family history of colorectal cancer between 2000 and 2009 inclusive. The aim was to assess and assign their risk of CRC according to BSG guidelines and subsequently identify the patients that developed CRC by following them up to the end of 2018. The approval authorized the use of the clinical genetics database to collect the relevant family history and genetic data, the endoscopic database for the follow-up of the patients and outcomes of colonoscopies, as well as the pathology database for polyp and cancer histopathology results. Since there was no patient contact for this study, ethics committee approval was not required.

3.1.2 Inclusion and Exclusion Criteria

This was a longitudinal study, with a retrospective patient cohort. Any individual referred to genetic counselling regarding FH of CRC from 2000-2009 was included in the study, regardless of age and gender. Patients were excluded if the available personal and FH information were more than 50% incomplete. Additional exclusion criteria include patients with a personal history of CRC or previous colectomy of any kind (total, subtotal etc). Furthermore, patients that tested negative for the known familial mutation were excluded in the study as they are considered to be at population risk.

3.1.3 Data collection and handling

Patients who met the inclusion criteria were identified and their FH information was collected from clinical genetics electronic records at Ninewells hospital and medical school, NHS Tayside. The data were extracted from clinical pedigrees, notes, FH questionnaires, as well as official correspondence to or from the clinical

genetics department (*See* **Table 9** for a comprehensive list of the clinical variables collected). If available, cancer data were extracted from the information services division Scotland. Any genetic testing on the patient or family member relating to a familial CRC condition was also identified through the genetics database. These included mutations identified through blood or specimen DNA testing.

Table 9 *List of clinical variables collected.*

Patient Information	Community Health Index Number (CHI Number) Date of Birth, Date of Death (if applicable) Pedigree number Date of first appointment in clinical genetics Age at first appointment in Clinical Genetics Any cancer diagnosis (type), Age of cancer diagnosis
History in first-degree relatives (FDRs)	<u>For mother and father:</u> Any cancer diagnosis(type), age of cancer diagnosis <u>For sister(s)/brother(s):</u> Total number of sisters/brothers, number of half-sisters/half-brothers Any cancer diagnosis (type), age of cancer diagnosis <u>For daughter(s)/son(s)</u> Total number of daughter(s)/son(s) Any cancer diagnosis (type), age of cancer diagnosis
History in second-degree relatives (SDR's)	Number of total SDRs Any cancer diagnosis (type), age of cancer diagnosis Whether SDR is paternal or maternal relative
Mutation testing	Mutation testing in patient (YES/NO) Relative with mutation testing (YES/NO) Type of mutation testing (blood/pathology specimen) Which mutations tested Mutation result (positive/negative) and if positive, which mutation
Endoscopic data	Year of endoscopy (colonoscopy) Outcome of endoscopy (polyp/no polyp/cancer) Risk of polyp (according to BSG criteria)- number, size Excision of polyp (yes/no/partly)
Pathology data	Polyp/cancer biopsy Polyp/cancer result of biopsy (adenoma/hyperplastic/other)
Surgical data	Any type of colectomy (yes/no) Year of colectomy

People within the cohort who were followed up by colonoscopy were identified through the endoscopy databases and Clinical Portal from years 2000 to 2018 inclusive within NHS Tayside. Data collected included the year of colonoscopy as well as colonoscopic findings, limited to normal, polyps and cancer. Information on polyps included the number of polyps and the size to assign a risk category (Low, Intermediate or High) according to BSG criteria, as well as if all, some or no polyps were extracted and retrieved. The histopathology of any polyps retrieved (Adenomas, Tubular adenomas, Tubulovillous adenomas, Hyperplastic polyp) and cancer pathology was identified through the pathology database.

All the patient identifiable information was stored in a secure disk on a Ninewells Hospital computer network in NHS Tayside. For safe handling of data out with this setting, community health index (CHI) number, date of birth and pedigree number were removed, and a study identification number was assigned to each patient. The pseudo-anonymised data were stored in a password protected file on a password protected laptop.

3.2 Assigning BSG Risk category

3.2.1 HNPCC

All people included in the study were assigned *de novo* into their corresponding risk categories according to BSG criteria based on FH information. All subcategories of FH that are included in each risk category can be seen in **Table 10**. Categories included low, moderate-low, moderate-high and high-risk. High-risk patients were subdivided into the ones with a confirmed identified mutation and the ones with high-risk FH but without an identifiable mutation. People who did not meet high risk FH criteria but had a confirmed mutation were allocated in high-risk category. A family member with two colorectal cancers or two different HNPCC associated cancers was counted as two individuals.

A panel of experts comprising of a consultant clinical geneticist and genetic counsellor discussed any patients that did not meet specific FH of BSG

categories. These patients were assigned to risk categories according to FH information and clinical judgement.

Table 10 *BSG risk categories and Family history criteria.* Information extracted from source 37.

BSG Risk Category	*Fulfilled criteria for each category*
High risk	• Confirmed mutation carrier • Fulfils Amsterdam Criteria • Untested FDR of proven mutation carrier • 1 FDR with MSI-H colorectal cancer AND IHC shows loss of MSH2, MSH6 or PMS2 expression.
Moderate- High risk	• Colorectal cancer in 3 FDR in first degree kinship, none <50 years • Colorectal cancer in 2 FDR in first degree kinship, mean age <60 years
Moderate-Low risk	• Colorectal cancer in 1 FDR <50 years • Colorectal cancer in 2 FDR in first degree kinship, mean age ≥ 60 years
Low risk	• Other Family History of Colorectal Cancer

3.2.2 Polyposis conditions

Patients with multiple polyps and suspected polyposis conditions were placed into their respective condition category. Patients with familial adenomatous polyposis (FAP) were dividing into at risk FAP if they were a member of an FAP family with no mutation identified or at high risk if they fulfilled the clinical criteria of FAP or were proven APC mutation carriers. Any FAP patients that underwent colectomy prior to referral were excluded from the study.

3.2.3 Polyps

For the identification of adenomatous polyps during surveillance, risk categories were assigned according to BSG adenoma surveillance. Low, Intermediate and High risk were assigned according to the number and size of adenomatous polyps as per BSG guidelines (*See* **Table 1**). Surveillance time may have been modified from that proposed by the guidelines according to polyps and FH.

3.3 Mutation carriers

3.3.1 HNPCC

People with HNPCC associated mutations within the cohort were identified. In cases where family history would have placed them into a low or a moderate-risk group, a confirmed mutation would mean that these patients fall into the high-risk category, as per guidelines. People with confirmed mutations that were assigned as variants of uncertain significance (VUS), according to American College of Medical Genetics and Genomics (ACMG) were assigned to a category according to FH criteria.

3.3.2 Polyposis

People who tested positive for mutations for FAP and attenuated familial adenomatous polyposis were assigned to a separate high-risk category. Patients that were FDR of a known mutation carrier were put into high risk category, even if they were not tested themselves. Similarly, patients were assigned to be at risk if there was clinical evidence (i.e. multiple polyps on colonoscopy) of FAP or AFAP even if a mutation was not found on testing. A few other mutations predisposing to polyposis conditions were identified in the cohort. These were assigned a high-risk, but no further analysis was performed as the statistical power would not be sufficient to reliably detect significant differences.

3.4 Polyps and cancer

People who were followed up through colonoscopic screening were identified as described in **Section 3.1.1**. Subsequently, those that developed polyps were also assigned a polyp risk category, according to BSG criteria (*See* **Section 3.2.3** and **Table 1**). Additionally, information was collected if polyps were excised or not. Analysis was performed on patients developing adenomatous polyps as these are regarded as cancer precursor lesions and they have a high risk of developing into an adenocarcinoma if left unexcised. The number of people in the cohort who also developed adenocarcinoma were identified as previously described and analysis was performed to determine the risk of developing adenocarcinoma in each risk category.

3.5 Statistical analysis

Relative Risk (RR), Odds Ratio (OR), Positive Predictive Value (PPV), Negative Predictive Value (NPV) and incidence were calculated, using the low-risk cohort as a reference group. SPSS statistics software and Microsoft Excel were used for data analyses. For analyses generating a significance value (*p*-value), $p \leq 0.05$ was used to determine significance.

3.5.1 Percentage 10-year absolute risk calculation

A % 10-year absolute risk for people within the cohort for each BSG risk category as well as HNPCC mutation carriers was calculated. This was performed for risk for the age groups 0-49 and above 50. The number of years follow-up between those age ranges was added to get the people years of follow-up. The % incidence of CRC per person year of follow up in both the age categories was calculated and multiplied by 10 to give the approximate 10-year risk for that age range. Similarly, the % incidence of adenomatous polyps per person per year of follow-up and the 10-year absolute risk for the age ranges was calculated. A % 5-year absolute risk was not calculated due to the limited cohort size.

3.5.2 Chi-square test, Fischer's exact test, independent t-test and one-way analysis of variance (ANOVA)

In order to calculate categorical variables such as cancer development (yes/no), Pearson's chi-squared test was used when numbers were sufficient or Fischer's exact test to report the asymptomatic two-sided p-value. The categorical variables analysed against cancer development is shown in **Table 11**.

To compare continuous variables (*See* **Table 11**), independent t-test was used. Levene's test was used to screen whether the data were normally distributed. Where Levene's test was significant, the reported *p*-value (2-sided) for the independent T-test does not assume equal variance of the continuous data. These variables were selected as broad descriptors for the family history of cancer.

To analyse the means of continuous variables with more than two groups for significance, an omnibus one-way analysis of variance (ANOVA) was performed.

Table 11 *Categorical and continuous variables for Chi-squared and t-test analysis respectively.*

Categorical variables analysed in Chi-Squared analysis	Continuous variables analysed in Independent T-test analysis
1 affected FDR or SDR with CRC cancer ≥ 2 affected FDR or SDR with CRC cancer ≥ 3 affected FDR or SDR with CRC cancer Parent affected with CRC Sibling affected with CRC Average age of relatives at diagnosis with CRC <50 Average age of relatives at CRC diagnosis <60	Proportion of FDRs and SDRs affected with CRC Average age of all relatives at CRC diagnosis Proportion of FDRs affected with CRC Average age of FDRs at CRC diagnosis Proportion of SDRs affected with CRC Average age of SDRs with CRC diagnosis
1 affected FDR or SDR with HNPCC associated cancer ≥ 2 affected FDR or SDR with HNPCC associated cancer ≥ 3 affected FDR or SDR with HNPCC associated cancer Parent affected with HNPCC associated cancer Sibling affected with HNPCC associated cancer Average age of relative at diagnosis with HNPCC cancer <50 Average age of relatives at HNPCC cancer diagnosis <60	Proportion of FDRs and SDRs affected with HNPCC cancer Average age of all relatives at HNPCC diagnosis Proportion of FDRs affected with HNPCC cancer Average age of FDRs at HNPCC diagnosis Proportion of SDRs affected with HNPCC cancer Average age of SDRs at HNPCC diagnosis
1 affected FDR or SDR with CRC or HNPCC cancer ≥ 2 affected FDR or SDR with CRC or HNPCC cancer ≥ 3 affected FDR or SDR with CRC or HNPCC cancer Parent affected with CRC or HNPCC cancer Sibling affected with CRC or HNPCC cancer Average age of relatives at diagnosis with CRC or HNPCC cancer <50 Average age of relatives at CRC or HNPCC diagnosis <60	Proportion of FDRs and SDRs affected with CRC and HNPCC Average age of all relatives at CRC and HNPCC diagnosis Proportion of FDRs affected with CRC and HNPCC Average age of FDRs at CRC and HNPCC diagnosis Proportion of SDRs affected with CRC and HNPCC Average age of SDRs at CRC and HNPCC diagnosis

3.5.3 Kaplan-Meier survival analysis

KM survival analysis was used to asses both adenocarcinoma and adenomatous polyp detection across the different BSG risk categories. Time was measured in number of years of follow-up for each risk category and the patients were censored at death or colectomy. Endpoints included CRC or first adenomatous polyp development depending on the KM curve and the end of the follow-up, the latest being the end of 2018. Separate analyses were performed to compare each BSG risk category for cancer or polyp development. Two separate analyses were conducted for the high-risk group, both including and excluding mutation carriers. For the survival curves looking at polyp detection, age-dependent analysis was also performed, for ages 0-49 and ≥50 years. KM survival curves were generated and presented for selected sets of results.

3.6 Sample size calculation

In order to address whether the study was adequately powered, a retrospective sample size calculation was performed on the categorical data: i.e. CRC incidence. To do so, the methodology from Jones, Carley and Harrison was used for studies reporting categorical data, i.e. Diagnosis of CRC.(182) The reason for performing the power calculation retrospectively was that when the study was initiated, the number of available patients was unknown and therefore not possible to ascertain power. The following assumptions were made when calculating the sample size:

- BSG guidelines do not effectively distinguish between medium and high-risk groups of familial CRC.
- Type 1 error is to be avoided at the conventional level of .05 ($p\alpha$).
- Type 2 error is to be avoided at the conventional level of .8 ($p\beta$).
- The clinically important difference to be detected is the difference in % absolute risk between the risk categories.

The following risk levels were used:

- Population risk was calculated using invasive colorectal cancer rates per 100,000 reported for the year of 2017; the data is available to download from ISD Scotland.(183) The lifetime rate per 100,000 per year of developing CRC is 69.6. This equates to 0.07% per person lifetime risk.
- For the moderate risk group, BSG guidelines find the risk of CRC to be between 1 in 6-10, depending on different studies.(37) 1 in 6 was used as the risk for the calculation so that the sample size needed to detect the smallest possible difference could be calculated.
- For high risk groups, there is 1 in 5-13 risk of developing CRC, depending on gender.(37) 1 in 5 was used as the risk for the calculation so that sample size needed to detect the smallest possible difference could be calculated.

Firstly, the standardised difference between the proportions of expected CRC was calculated:

$$\text{Standardised difference} = p1\text{-}p2/\sqrt{(P\ (1-P)\ )}$$

p1 = risk of colorectal cancer in higher risk group

p2 = risk of colorectal cancer in lower risk group

$$P = (p1 + p2)/2$$

To work out the required sample size, a standardised risk is used and applied to the nomogram below shown in **Figure 7.** A line is drawn from the calculated standard difference across to the designated pβ and by looking at the pα level of 0.05, the required sample size can be deduced.

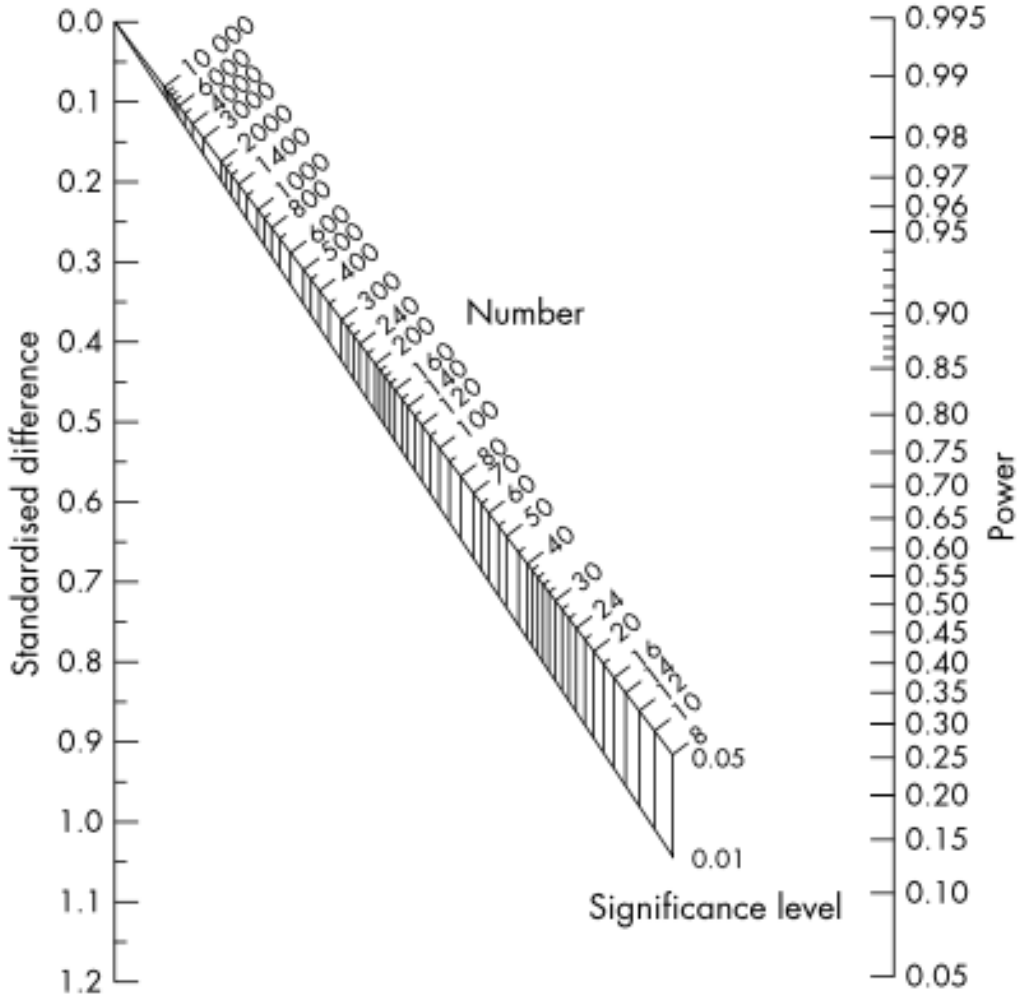


Figure 7 *Nomogram for the calculation of sample size.* Figure taken from source 182.

The results of the retrospective power calculations are as follows:

3.6.1 Sample size required to detect clinically significant difference between population and moderate risk group

$$\text{Standardised difference} = p_1 - p_2 / \sqrt{P(1-P)}$$

p_1=minimum risk of CRC in moderate risk group = 0.17 or 17%
p_2 = population risk of CRC=0.07 or 7%

$$P = (p_1 + p_2)/2$$

$$P = (0.17 + 0.07)/2 = 0.12$$

$$\text{Standardised difference} = (0.17 - 0.07)/\sqrt{0.12\ (1 - 0.12)}$$

$$\text{Standardised difference} = 0.1/\sqrt{0.1056}$$

$$\text{Standardised difference} = 0.1/0.325$$

$$\text{Standardised difference} = 0.308$$

Using the above nomogram, a standardised difference of 0.308 with a pβ of 0.8 and pα of 0.05, a sample size of 300 subjects is required to adequately power the study. There were 288 low and 316 moderate-risk subjects (604 in total) in the study cohort, meaning that the study is adequately powered to detect a clinical difference between low and moderate risk groups.

3.6.2 Sample size required to detect clinically significant difference between population and high-risk group

$$\text{Standardised difference} = p_1 - p_2 / \sqrt{(P\,(1-P))}$$

p_1=minimum risk of CRC in high risk group = 0.2 or 20%
p_2 = population risk of CRC=0.07 or 7%

$$P = (p1+p2)/2$$

$$P = (0.2 + 0.07)/2 = 0.135$$

$$\text{Standardised difference} = (0.2 - 0.07)/\sqrt{0.135\,(1 - 0.135)}$$

$$\text{Standardised difference} = 0.13/\sqrt{0.117}$$

$$\text{Standardised difference} = 0.13/0.342$$

$$\text{Standardised difference} = 0.380$$

Using the above nomogram, a standardised difference of 0.380 with a pβ of 0.8 and pα of 0.05, a sample size of 220 subjects is required to adequately power the study. There were 288 low and 124 high risk subjects (412 in total) in the study cohort. This means that the study is adequately powered to identify a clinical difference between low and high-risk subjects.

3.6.3 Sample size required to detect clinically significant difference between moderate and high-risk group

$$\text{Standardised difference} = p_1 - p_2 / \sqrt{(P\,(1-P))}$$

p_1= minimum risk of CRC in high risk group = 0.2 or 20%
p_2 = minimum risk of CRC in moderate risk group = 0.17 or 17%

$$P = (p1+p2)/2$$

$$P = (0.2 + 0.17)/2 = 0.185$$

$$\text{Standardised difference} = (0.2 - 0.17)/\sqrt{0.185\,(1 - 0.185)}$$

$$\text{Standardised difference} = 0.03/\sqrt{0.151}$$

$$\text{Standardised difference} = 0.03/0.389$$

$$\text{Standardised difference} = 0.077$$

Using the above nomogram, a standardised difference of 0.077 with a pβ of 0.8 and pα of 0.05, a sample size of 4000 subjects is required to adequately power the study. There were 124 high and 316 moderate risk subjects (440 in total) in the study cohort. This means that the study is substantially under-powered to identify a clinical difference between medium and high-risk subjects.

Chapter 4: Results

4.1 Descriptive Statistics

4.1.1 Cohort Characteristics

For this study, 1120 patients in total presented in the clinical genetics department between 2000-2009 with a family history of colorectal cancer. From those, 183 were excluded because of previous cancer or a previous colectomy and 3 more records were excluded as they did not have any FH of CRC. 22 records were excluded due to incomplete FH data. A further 132 patients were excluded because they had a first degree relative (FDR) with a known mutation which predisposes to an increased risk of CRC, but they themselves tested negative of that mutation. In total, 780 patients met the criteria for inclusion. From those patients, 52 were referred for polyposis conditions; for the complete set of results for the polyposis patients, *See* **Section 4.6.**

There was a total of 728 patients with a non-polyposis FH of CRC. The age of presentation in clinic ranged from 0 years 11 months to 84 years. The mean age of presentation was 44.948 (standard deviation (SD) ±12.2472) years and the median age was 44 years. All patients included had no personal history of CRC, however, there were 30 incidences of other cancers. The total patient years follow-up for cancer (excluding polyposis patients) was 5561. The mean number of years of follow up was 12.696 (SD ±2.8092) with a median of 13 years.

4.1.2 Family history structures and cancer history

Table 12 summaries the incidence of each type of cancer in the patient cohort. It is worth mentioning that none of the patients that had a personal history of other cancer developed CRC later on.

Table 12 *Incidence of cancer in cohort.*

TYPE OF CANCER	INCIDENCE	
	Frequency	**% cohort**
BASAL CELL CARCINOMA	3	0.41%
MELANOMA	3	0.41%
CERVIX	1	0.14%
UTERUS	5	0.69%
BREAST	12	1.65%
BONE	1	0.14%
OVARY	2	0.27%
LUNG	1	0.14%
HAEMATOLOGICAL	2	0.27%
TOTAL	**30**	**4.12 %**

The FH structure from clinical notes is recorded in **Table 13**.

Table 13 *Family history structure of cohort.*

	Mean	**Range**	
		Minimum	**Maximum**
Sisters (Including half-sisters)	1.22	0	10
Brothers (Including Half-brothers)	1.12	0	9
Total number of siblings	2.35	0	10
Daughters	0.65	0	4
Sons	0.64	0	4
Total number of Children	1.29	0	6
Second Degree relatives	10.82	4	39

Information regarding the incidence and percentage of the cohort who reported cancer in various family members is described in **Table 14.** The table is divided in incidence and percentage of family members presenting with CRC, other HNPCC associated cancer and other cancers. The most common relative reported with CRC was the mother (34.75%) and the most common affected relative with other HNPCC associated cancer was the father (7.42%). From the patient cohort, there were 587 reports of CRC in a second degree relative.

Table 14 *Cancer incidence in relatives of the cohort.*

	Incidence in relative amongst cohort (N=728)	
	N	% of cohort
Colorectal cancer incidence		
Mother	253	34.75%
Father	240	32.97%
Sibling(s)	192	26.37%
Child(ren)	6	0.82%
SDR	587	80.63%
Other HNPCC associated cancer incidence*		
Mother	54	7.42%
Father	58	7.97%
Sibling(s)	59	8.10%
Child(ren)	5	0.69%
SDR	330	45.33%
Other cancer incidence		
Mother	134	18.41%
Father	62	8.52%
Sibling(s)	86	11.81%
Child(ren)	9	1.24%
SDR	632	86.81
*Other HNPCC cancer incidence included endometrial (females), gastric, ovarian (females), small bowel, bladder, brain, kidney, biliary tract, liver, gallbladder and pancreatic cancer.		

Table 15 Shows family data regarding age of diagnosis of CRC and other HNPCC associated cancers in the whole cohort.

Table 15 *Mean age of relatives at cancer diagnosis.*

	Mean (±SD)	Median	Range	
			Minimum	Maximum
Age of total FDRs and SDRs with CRC diagnosis	59.635(±13.9764)	60	19	100
Age of total FDRs and SDRs with HNPCC cancer diagnosis	59.013(±14.8450)	60	11	95
Age of FDRs with CRC diagnosis	57.110(±13.0720)	57	19	93
Age of FDRs with HNPCC cancer diagnosis	54.855(±15.0130)	54	11	87
Age of Mother with CRC diagnosis	59.780(±12.9384)	60	23	93
Age of Mother with HNPCC cancer diagnosis	58.137(±12.235)	55	27	87
Age of Father with CRC diagnosis	59.177(±12.1805)	59	20	87
Age of Father with HNPCC cancer diagnosis	61.782(±11.5706)	64	39	83
Age of Sibling(s) with CRC diagnosis	51.654(±12.1038)	50	28	78
Age of Sibling(s) with HNPCC cancer diagnosis	47.691(±13.9660)	50	11	76
Age of Child(ren) with CRC diagnosis	35.167(±9.4060)	36.5	19	46
Age of Child(ren) with HNPCC cancer diagnosis	24(±13.5499)	16	13	48
Age of Total SDRs with CRC diagnosis	62.801(±14.0262)	63	20	100
Age of Total SDRs with HNPCC cancer diagnosis	61.487(±14.1768)	62	20	95

4.1.3 BSG Risk Categories

After assigning each patient to the corresponding BSG category according to FH and mutation status, there were 288 (39.56%) low-risk patients, 316 (43.41%) medium-risk and 124 (17.03%) high-risk patients. The mean age at presentation in these groups was 44.51 (SD±11.3022), 45.61 (SD±12.0352) and 44.26 (SD±12.5681) respectively. One-way analysis of variance (ANOVA) demonstrated no significant difference when comparing the age of assessment amongst the three groups; F (2,725) = 0.836, p=0.434. **Table 16** breaks down each FH subtypes that belong to each risk category. Of those in the moderate risk category, the majority (13.87%) had one FDR with CRC under the age of 50, which falls in the low-moderate sub-category. From those in the high-risk group, the majority (11.95%) were at risk of being an MMR gene carrier that fulfilled the Amsterdam Criteria. It is worth noting that 71 patients did not fully meet the BSG FH criteria and were assigned at appropriate risk by a panel of experts according to FH and clinical judgement.

Table 16 *Percentage of cohort that fit into each family history category according to BSG guidelines.*

	Frequency	% of cohort
Moderate risk		
Colorectal cancer in 1 FDR <50 years	101	13.87%
2 FDR ≥60 years in first degree kinship	82	11.26%
Colorectal cancer in 3 FDR in first degree kinship, none <50 years	63	8.65%
Colorectal cancer in 2 FDR in first degree kinship, mean age <60 years	70	9.62%
	316	43.4%
High risk		
Confirmed mutation carrier	31	4.26%
At risk HNPCC MMR carrier that Fulfils Amsterdam Criteria	87	11.95%
Untested FDR of proven mutation carrier	5	0.69%
1 FDR with MSI-H colorectal cancer AND IHC shows loss of MSH2, MSH6 or PMS2 expression.	1	0.14%
	124	17.0%

4.1.4 Mutation testing

From the cohort, 54 patients were tested and 31 of them were found to have a mutation that increased their risk of CRC and 2 had a variant of uncertain significance (VUS).

There were 142 patients that had a relative tested without a mutation identified; however, 6 of them had a VUS. The complete set of results in terms of patients and their relatives that was tested is found in **Table 17**. **Table 18** shows all the mutations identified in the cohort.

Table 17 *Percentage of cohort and their relatives that were tested for mutations in each risk category.*

	Patient		**Relative**	
	N	**% Cohort**	***N***	**% Cohort**
Low risk	10	1.37%	30	4.12%
Medium Risk	6	0.82%	80	10.99%
High risk	38	5.22%	32	4.40%
Total	54	7.4%	142	19.5%

Table 18 *All Lynch syndrome mutations identified in the cohort and their frequency.*

Mutations	**Frequency**
MLH1	
MLH1 381-1 G>C	2
MLH1 G67E mutation	1
MLH1/A681T exon 18	1
MLH1 c.1190delT	2
MLH1 c.00G>A p.gly67Glu	1
MLH1 pa681T exon 18	4
MLH1/c.1017delC	2
MLH1/R265C	1
MLH1/E102D heterozygous	1
MLH1 c.117-1G>C	1
MLH1 c.473delA	1
MSH2	
MSH2 c.628delAT	4
MSH2 del exons 1-6	2
MSH2c388_389delCA	1
MSH6	
MSH6 c.3261dupC	1
MSH6 c.3518_3519insA	1
MSH6 del Exons 5-6	1
Other	
PMS2 exon 7 deletion	1
PMS2 c.137G>T	3
Total	**31**

4.1.5 Number of years of follow-up

The number of people years of follow up before CRC development for each BSG risk category is shown in **Table 19**. Patients were also divided according to age of presentation.

Table 19 *Follow-up years for each risk category according to age group for colorectal cancer development.*

		0-49 years	≥ 50 years	Total
Low risk		943	433	1376
Moderate risk		1825	1027	2852
High risk	Non-Mutation Carriers	561	433	994
	Mutation Carriers	248	91	339
Total		3577	1984	5561

Table 20 shows the number of people years of follow-up before polyp detection for each BSG risk category.

Table 20 *Follow-up years for each risk category according to age group for polyp detection.*

		0-49 years	≥ 50 years	Total
Low risk		885	413	1298
Moderate risk		1763	892	2655
High risk	Non-Mutation Carriers	459	380	839
	Mutation Carriers	228	91	319
Total		3335	1776	5111

4.1.6 Cancer Development in cohort

In total, 8 people developed colorectal adenocarcinoma from a cohort of 728. Of the eight cancers developed, two (25%) were in the low-risk category, three (37.5%) in the moderate-risk and three (37.5%) in the high-risk, including two mutation carriers. There was a significantly higher likelihood of cancer development in the mutation risk group compared to the low risk group (Fischer's exact p=0.0485) but no other significant difference between categories (*See* **Table 21** for complete set of results). The patients first presented to clinical genetics with a mean age of 54.710 (SD ± 17.2818), with median age of 62 and an age range of 28-79. The mean age of CRC diagnosis across all three categories was 63.460 years (SD ±14.0337) and the median age was 65 years, ranging from 38-79. The mean time from first presentation at clinical genetics to the development of CRC was 8.750 years (SD ± 5.8041), ranging from 0 to 16 years.

Table 21 *Significance of colorectal cancer development between categories.*

Risk categories compared	*p*-value (Fischer's exact)
Low & Moderate	1.0000
Low & High (including mutation carriers)	0.1631
Low & High (excluding mutation carriers)	0.5692
Low & Mutation carriers	**0.0485**
Low & Moderate *plus* High (Including mutation carriers)	0.4892
Moderate & High (including mutation carriers)	0.3570
Moderate & High (excluding mutation carriers)	1.0000
Moderate & Mutation carriers	0.0652

4.1.7 Polyp Detection in cohort

In total, there were 65 patients with adenomatous polyps detected. Of those people who had adenomatous polyps, 11 (16.9%) were in the low-risk group, 31 (47.7%) in the moderate-risk group and 23 (35.4%) in the high-risk group, including five in the mutation carrier group. Fischer's exact tests comparing the likelihood of polyp detection between categories is presented in **Table 22**. The low-risk group has a significantly lower chance of having a polyp detected compared to the moderate and high-risk groups, both including and excluding the mutation carriers. Furthermore, there was a significantly higher chance of polyp detection in the high-risk groups, including and excluding mutation carriers compared to the moderate risk group (p=0.0153 and p=0.0177 respectively). However, there was no significant difference in the likelihood of polyp detection between the moderate and mutation carrier groups (p=0.3475). Those patients first presented in the genetics clinic with a mean age of 47.785 years (SD ±10.3688), median age 47 and age range 28-68. The age of first adenoma detection ranged from 31-74. The mean age of first adenomatous polyp detection was 53.416 (SD ± 11.2193), median age was 54. The mean time from first appointment in clinical genetics to adenoma development was 5.631 (SD ± 4.8216) years and it ranged from 0-15 years.

Table 22 *Significance of polyp detection between categories.*

Risk categories compared	*p*-value (Fischer's exact)
Low & Moderate	**0.0038**
Low & High (including mutation carriers)	**0.00001**
Low & High (excluding mutation carriers)	**0.00001**
Low & Mutation carriers	**0.0126**
Low & Moderate *plus* High (Including mutation carriers)	**0.0001**
Moderate & High (including mutation carriers)	**0.0153**
Moderate & High (excluding mutation carriers)	**0.0177**
Moderate & Mutation carriers	0.3475

4.1.8 Cohort Summary

Figure 8 summarises the data collection process and basic descriptive characteristics of the cohort.

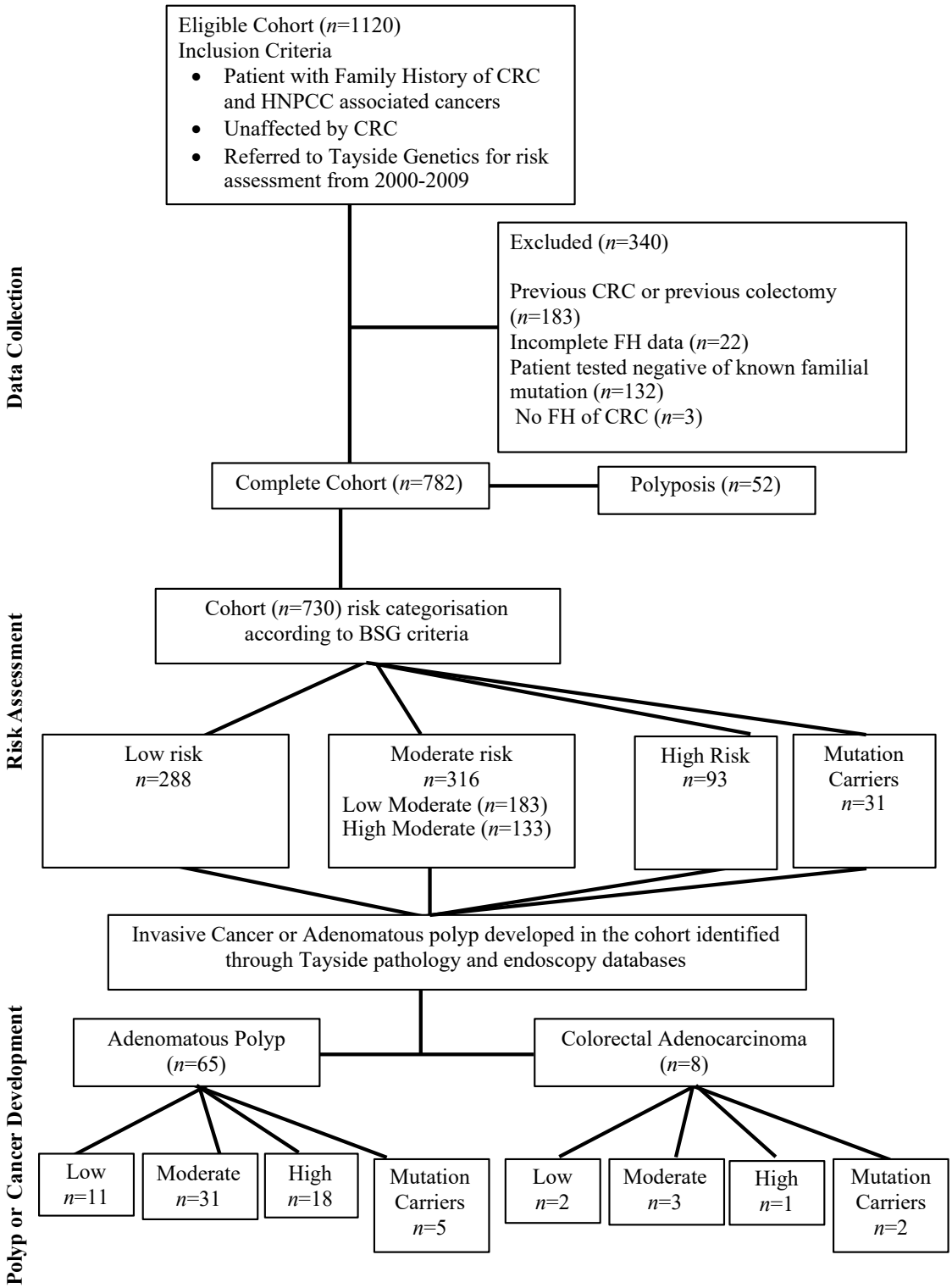


Figure 8 *Summary of data collection process and cohort characteristics.*

4.2 Risk analysis

4.2.1 CRC development by BSG category

Mean age of cancer diagnosis for each group is shown in **Table 23** and plotted in **Figure 9.** One-way ANOVA demonstrated no significant difference between age of presentation of CRC across the three risk categories, including and excluding mutation carriers (F(2,5)=0.0247, *p*=0.9757 and F(2,3)=0.5673, *p*=0.6180*);* however the number of cancers developed and thus analysed was small in each category, therefore the analysis is substantially underpowered.

Table 23 *Mean age of CRC development in each category.*

	Mean (±SD)	Median	Range
Low risk	61.161(±15.1407)	61	46-76
Moderate Risk	63.753(±2.7387)	64	61-67
High Risk (including mutation carriers)	64.720(±18.9088)	77	38-79
High Risk (excluding mutation carriers)	77(±0)	77	77
Mutation carriers only	58.500(±20.500)	59	38-79

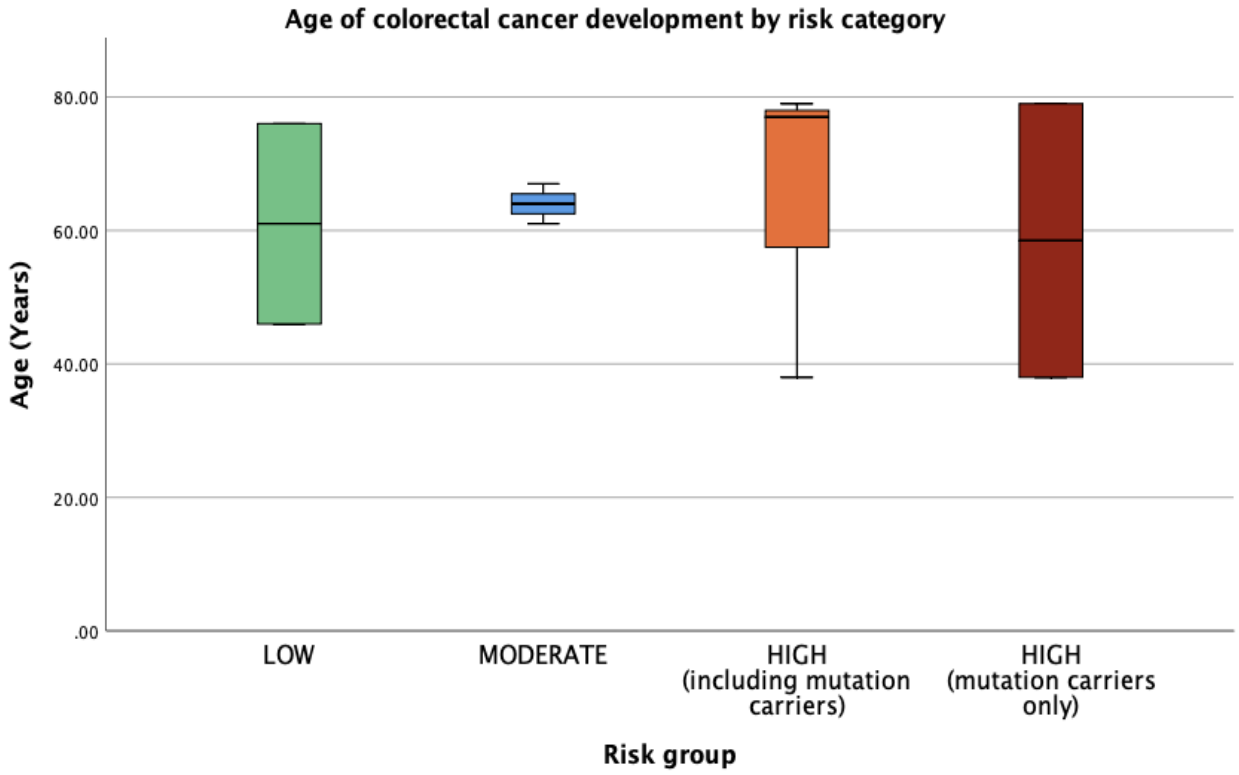


Figure 9 *Box plot for age of colorectal cancer development in each risk category.*

4.2.2 Mean age of adenoma detection t by BSG category

Mean age of polyp diagnosis for each group is described in **Table 24** and plotted in **Figure 10**. One-way ANOVA demonstrated a significant difference between the age of first polyp development in the three risk groups when mutation carriers were included, ($F(2,62)=5.2262$, $p=0.0080$) and when excluded ($F(2,57)=3.3300$, $p=0.0429$). Three post-hoc Bonferroni-corrected independent-samples t-tests were conducted to examine the relationship between age and polyp development. These showed that the age of adenomatous polyp detection is significantly lower in the high-risk group relative to the moderate-risk group, including and excluding mutation carriers ($p=0.0017$ and $p=0.0103$ respectively). However, there was not a significant difference in the age of polyp detection between low and moderate-risk groups ($p=0.2613$) and low and high-risk groups ($p=0.2723$ including mutation carriers and $p=0.4696$ excluding mutation carriers).

Table 24 *Mean age of polyp detection in each category.*

	Mean (±SD)	Median	Range
Low risk	52.831(±11.3362)	50	34-74
Moderate Risk	57.554(±10.3938)	58	33-74
High Risk (including mutation carriers)	48.118(±9.8660)	51	31-65
High Risk (excluding mutation carriers)	49.719(±9.0740)	52	31-65
Mutation carriers only	42.351(±10.4331)	37	32-59

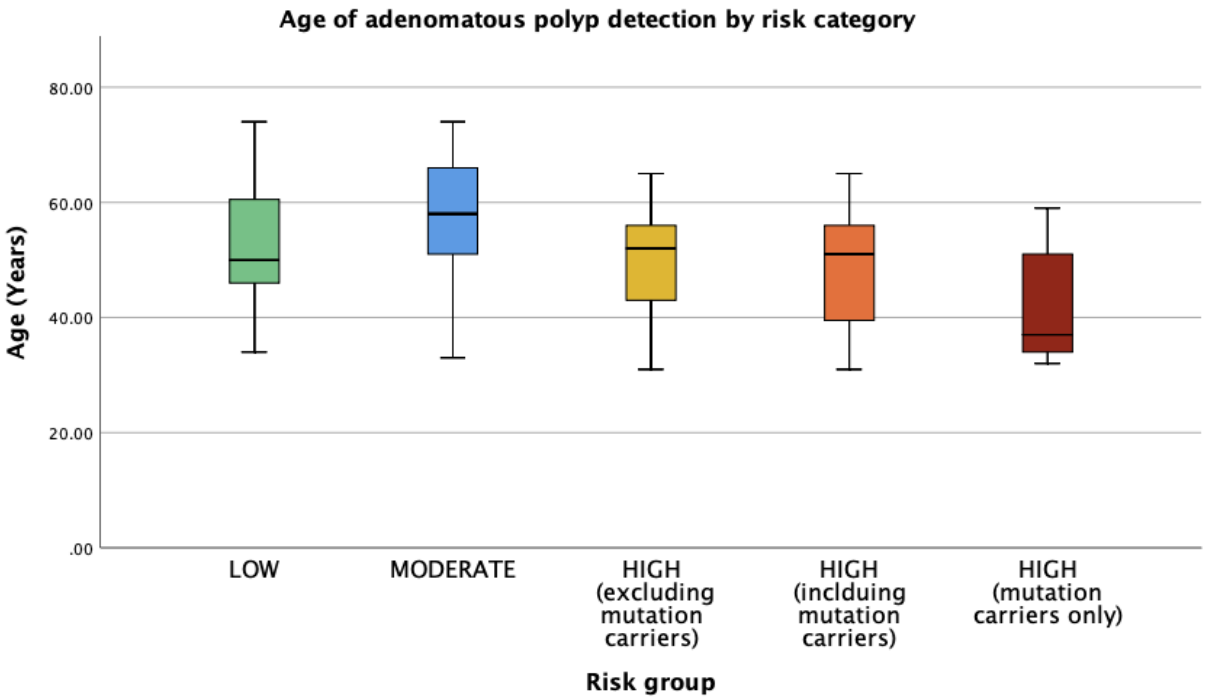


Figure 10 *Box plot for age of adenomatous polyp detection in each risk category.*

4.2.3 Independent T-test analysis

In order to compare continuous variables with CRC development, independent T-test analysis was performed. The full results of the analysis can be found in **Table 25.** It should be noted that the significance values presented are not corrected for multiple comparisons.

4.2.4. Pearson Chi-square/Fischer's exact tests

The full results of Pearson Chi-Square to compare the collected categorical variables with CRC can be found in **Table 26**. It should be noted that the significance values presented are not corrected for multiple comparisons.

Table 25 *Independent t-test analysis of continuous family history variables in relation to cancer development, including and excluding mutation carriers.*

	Including Mutation Carriers				**Excluding Mutation Carriers**			
	Cancer Group (N=8)	Non-Cancer Group (N=720)	Independent t-test		Cancer Group (N=6)	Non-Cancer Group(N=691)	Independent t-test	
Variables	M (±SD)	M (±SD)	F	*p*	M (±SD)	M(±SD)	F	*p*
Proportion of FDRs and SDRs affected with CRC	.129(±.0887)	.126(±.0884)	.004	.918	.120(±.1032)	.127(±.0888)	.275	.857
Average age of all relatives at CRC diagnosis	60.217(±13.5795)	58.628(±11.7103)	.228	.741	67.875(±8.5281)	59.086(±11.5334)	.788	.129
Proportion of FDRs affected with CRC	.292(±.3611)	.240(±.2054)	5.739	.696	.353(±.4001)	.242(±.2055)	8.773	.529
Average age of FDRs at CRC diagnosis	67.4000(±13.6675)	56.285(±12.3884)	.000	**.047**	72.500(±8.6989)	56.711(±12.2087)	.473	**.010**
Proportion of SDRs affected with CRC	.088(±.0819)	.084(±.1024)	.493	.913	.063(±.0758)	.085(±.1049)	.667	.603
Average age of SDRs at CRC diagnosis	54.133(±14.4714)	63.337(±12.2442)	.003	.097	60.667(±16.0728)	63.850(±12.0872)	.382	.651
Proportion of FDRs and SDRs affected with HNPCC cancer	.054(±.0563)	.045(±.0613)	.608	.689	.062(±.0615)	.044(±.0605)	.275	.490
Average age of all relatives at HNPCC diagnosis	68.250(±10.8819)	59.076(±14.4995)	.812	.209	72.833(±7.1822)	59.739(±14.1624)	1.174	.111
Proportion of FDRs affected with HNPCC cancer	.014(±.0393)	.055(±.1380)	3.127	.402	.00(±.00)	.053(±.1369)	4.140	**.000**
Average age of FDRs at HNPCC diagnosis	67.000(±0)	55.6509(±15.6100)	-	.470	-	-	-	-
Proportion of SDRs affected with HNPCC cancer	.066(±.0751)	.043(±.0720)	.019	.371	.081(±.0816)	.043(±.0719)	.028	.197
Average age of SDRs at HNPCC diagnosis	65.125(±16.4943)	60.805(±13.5949)	.038	.541	72.833(±7.1822)	61.406(±13.5485)	1.093	.147
Proportion of FDRs and SDRs affected with CRC and HNPCC	.183(±.1379)	.171(±.1015)	.873	.741	.182(±.1618)	.171(±.1031)	2.560	.804
Average age of all relatives at CRC and HNPCC diagnosis	62.245(±13.5664)	58.327(±11.0925)	.527	.390	70.225(±7.1369)	58.919(±10.7251)	.875	**.036**
Proportion of FDRs affected with CRC and HNPCC	.306(±.3601)	.295(±2237)	3.784	.890	.353(±.4000)	.295(±.2249)	6.153	.739
Average age of FDRs at CRC and HNPCC diagnosis	68.667(±11.4115)	56.205(±12.3959)	.044	**.025**	72.500(±8.6987)	56.716(±12.0969)	.314	**.009**
Proportion of SDRs affected with CRC and HNPCC	.155(±.1197)	.128(±.1234)	.017	.539	.144(±.1394)	.128(±.1247)	.248	.758
Average age of SDRs at CRC and HNPCC diagnosis	61.913(±16.0685)	62.597(±11.6592)	.928	.890	70.833(±10.6197)	63.218(±11.4375)	.002	.185

	Including Mutation Carriers					Excluding Mutation Carriers				
	Cancer Group	Non-Cancer Group				Cancer Group	Non-Cancer			
Variables	Count (expected)	Count (expected)	df	Chi^2	*p*	Count (Expected)	Count (Expected)	df	Chi^2	*p*
1 affected FDR or SDR with CRC	0(3)	271(286)	1	4.797	**.029**	0(2.2)	261(258.8)	1	3.623	.057
≥ 2 affected FDR or SDR with CRC	6(4.3)	381(382.7)	1	1.550	.213	4(3.2)	368(368.8)	1	.430	.512
≥ 3 affected FDR or SDR with CRC	3(1.7)	152(153.3)	1	1.268	.260	2(1.3)	147(147.7)	1	.515	.473
Parent affected with CRC	5(5)	448(448)	1	.000	.987	4(3.8)	434(434.2)	1	.038	.846
Sibling affected with CRC	2(1.9)	168(168.1)	1	.012	.912	2(1.4)	160(160.6)	1	.345	.557
Average age of relatives at CRC diagnosis <50	2(1.2)	131(131.8)	1	.576	.448	0(.8)	117(116.2)	1	.944	.311
Average age of relatives at CRC diagnosis <60	3(3)	313(313)	1	.001	.978	1(1.9)	293(292.1)	1	.834	.361
1 affected FDR or SDR with HNPCC associated cancer	200(200.8)	3(2.2)	1	.372	.542	3(1.7)	197(198.3)	1	1.343	.247
≥ 2 affected FDR or SDR with HNPCC associated cancer	2(1.4)	122(122.6)	1	.363	.547	1(1)	111(111)	1	.002	.968
≥ 3 affected FDR or SDR with HNPCC associated cancer	0(.3)	24(23.7)	1	.276	.599	0(.2)	21(20.8)	1	.188	.665
Parent affected with HNPCC associated cancer	0(1.1)	104(102.9)	1	1.348	.246	0(.8)	97(96.2)	1	.978	.323
Sibling affected with HNPCC associated cancer	1(.4)	32(32.6)	1	1.186	.276	0(.2)	29(28.8)	1	.263	.608
Average age of relative at HNPCC cancer diagnosis <50	0(.9)	69(68.1)	1	1.195	.274	0(.7)	63(62.3)	1	.845	.358
Average age of relatives at HNPCC cancer diagnosis <60	1(1.8)	137(1362.)	1	.690	.406	0(1.3)	125(123.7)	1	2.311	.128
1 affected FDR or SDR with CRC or HNPCC cancer	0(1.8)	161(159.2)	1	2.297	.130	0(1.4)	158(156.6)	1	1.774	.183
≥ 2 affected FDR or SDR with CRC or HNPCC cancer	6(5.9)	530(530.1)	1	.008	.529	4(4.4)	504(503.6)	1	.118	.731
≥ 3 affected FDR or SDR with CRC or HNPCC cancer	5(3.5)	310(311.5)	1	1.219	.270	4(2.6)	300(301.4)	1	1.308	.253
Parent affected with CRC or HNPCC cancer	5(5.8)	526(525.2)	1	.447	.504	4(4.4)	506(505.6)	1	.130	.718
Sibling affected with CRC or HNPCC cancer	3(2.1)	189(189.9)	1	.516	.473	2(1.6)	179(179.4)	1	.171	.679
Average age of relatives at CRC or HNPCC diagnosis <50	2(1.1)	124(124.9)	1	.904	.342	0(.6)	106(105.4)	1	.776	.378
Average age of relatives at CRC or HNPCC diagnosis <60	2(3)	344(343)	1	.708	.400	0(1.9)	318(316.1)	1	3.786	.052

Table 26 *Chi^2/Fischer's exact tests for categorical family history variables in relation to colorectal cancer development, including and excluding mutation carriers.*

4.2.5 Frequency and percentage 10-year absolute risk of CRC

The frequency of CRC diagnosis in the cohort by BSG risk category and age range is seen in **Table 27,** with mutation carriers shown both separately and included in the high-risk group. The % 10-year absolute risk (AR), based on people years of follow up in the cohort is shown for ages 0-49 and ≥50. The % 10-year AR was calculated as below:

$$\frac{number\ of\ cancers\ in\ age\ category}{number\ of\ years\ of\ follow-up\ in\ that\ age\ category}\ X\ 10\ X\ 100$$

e.g. % 10-year AR for the low risk group, between ages 0-49:

$$\frac{1}{943}\ X\ 10\ X\ 100 = 1.06\%$$

Across all age ranges, the fewest cancer diagnoses occurred in the low-risk group. The lowest % 10-year AR was in the 0-49 age-group, in the low-risk category (1.06%). The highest absolute risk for CRC development for the ages 0-49 was in the mutation carrier group (4.03%). The highest %10-year AR in ages ≥50 was again in the mutation carrier group (10.99%) and in the high-risk group (3.82%) when combined with the mutation carriers.

Table 27 *Percentage 10-year absolute risk of colorectal cancer development for each risk category and age group, including and excluding mutation carriers.*

	Mutation carriers separate				**Mutation carriers included**			
	N	**Number of CRC (% 10-year absolute risk (95% CI))**			**N**	**Number of CRC (% 10-year absolute risk (95% CI))**		
		Age range (Years)				**Age range (Years)**		
		Overall	**0-49**	**≥50**		**Overall**	**0-49**	**≥50**
Low Risk	288	2	1 1.06% (0.468-1.652)	1 2.31% (1.593-3.026)	288	2	1 1.06% (0.468-1.652)	1 2.31% (1.593-3.026)
Moderate Risk	316	3	0	3 2.92% (2.201-3.640)	316	3	0	3 2.92% (2.201-3.640)
High Risk	93	1	0	1 2.31% (1.372-3.292)	124	3	1 1.24% (0.625-1.847)	2 3.82% (2.808-4.829)
Mutation Carriers	31	2	1 4.03% (2.714-5.346)	1 10.99% (6.728-10.250)				
Total	728	8	2	6	728	8	2	6

4.2.6 Frequency and percentage 10-year absolute risk of adenomatous polyp detection

Table 28 Shows the % 10-year risk of polyp detection in the cohort by BSG risk category and age range. Across all age ranges, the highest polyp detection was in the high-risk group. The % 10-year AR of polyp detection in the 0-49 age group was the highest in the high-risk group (28.32% and 26.20%, excluding and including mutation carriers respectively), and the lowest in the moderate-risk group (7.37%). For ages ≥50, the % 10-year AR was the highest in the moderate-risk group (20.18%).

Table 28 *Percentage 10-year absolute risk of adenomatous polyp detection for each risk category and age group, including and excluding mutation carriers.*

	Mutation carriers separate				**Mutation carriers included**			
	N	**Number of adenomatous polyps (% 10-year absolute risk (95% CI))**			**N**	**Number of adenomatous polyps (% 10-year absolute risk (95% CI))**		
		Age range (Years)				**Age range (Years)**		
		Overall	**0-49**	**≥50**		**Overall**	**0-49**	**≥50**
Low Risk	288	11	8 9.04% (8.177-9.904)	3 7.26% (6.110-8.418)	288	11	8 9.04% (8.177-9.904)	3 7.26% (6.110-8.418)
Moderate Risk	316	31	13 7.37% (6.804-7.944)	18 20.18% (19.166-21.193)	316	31	13 7.37% (6.804-7.944)	18 20.18% (19.166-21.193)
High Risk	93	18	13 28.32% (26.979-29.661)	5 13.16% (11.500-14.620)	124	23	18 26.20% (25.143-27.257)	5 10.62% (6.605-14.635)
Mutation Carriers	31	5	5 21.93% (20.096-23.764)	0				
Total	728	65	39	26	728	65	39	26

4.2.7 Relative risks and odds ratios for BSG risk categories

Table 29 shows the RR and OR for the moderate and high-risk groups in comparison to the low risk group for the development of CRC. The mutation carriers are shown both as separate and included in the high-risk group. The RR and OR associated with both moderate and high-risk groups is also shown. The only significant finding is the higher RR and OR of CRC in mutation carriers compared to the low-risk group.

Table 29 Relative risks and Odds rations for colorectal cancer development in each category compared to the low-risk group.

	Relative risk (95% CI)	**Odds ratio (95% CI)**
Moderate risk	1.367 (0.2301-8.1236)	1.371(0.2274-8.2619)
High risk (including mutation carriers)	3.484 (0.5894- 20.5928)	3.546(0.5850-21.4876)
High risk (excluding mutation carriers)	1.548 (0.1420-16.8830)	1.554 (0.1393-17.3401)
Mutation carriers	9.290 (1.3557-63.6653)	9.862(1.3389-72.6439)
Moderate or high	1.964 (0.3991-9.6622)	1.977 (0.3962-9.8635)

The RR and OR for the moderate or high-risk groups for the detection of adenomatous polyps is shown in **Table 30**. The highest RR and OR for polyp detection is in the high-risk group excluding the mutation carriers, followed by the high-risk group including the mutation carriers. The lowest RR and OR of polyp detection is in the moderate risk group. These values are significant, confirming that there is a higher RR and OR of detecting adenomatous polyps in the high-risk groups compared to the low-risk group.

Table 30 Relative risks and Odds rations for adenomatous polyp detection in each category compared to the low-risk group.

	Relative risk (95% CI)	**Odds ratio (95% CI)**
Moderate risk	2.569 (1.3155-5.0148)	2.739 (1.3501-5.5571)
High risk (including mutation carriers)	4.856 (2.4432-9.6529)	5.735 (2.6986-12.1857)
High risk (excluding mutation carriers)	5.067 (2.4845-10.3355)	6.044 (2.7366-13.3473)
Mutation carriers	4.223 (1.5690-11.3659)	4.8423 (1.5628-15.0063)
Moderate or high	3.213 (1.7094-6.0399)	3.5223 (1.8090-6.8603)

4.2.8 Sensitivity, specificity, positive and negative predictive values of BSG risk categories

The sensitivity, specificity, positive and negative predictive values (PPV, NPV) of the risk categories are shown in **Table 31** for CRC development. These were calculated for the high-risk group, as well as the moderate and high-risk groups combined, compared with the low and moderate-risk groups and the low risk-group respectively. The sensitivity and specificity for the high-risk group was 2.42% and 99.17% respectively. Sensitivity was lower for the moderate and high-risk groups combined (1.36%) but the specificity was greater (99.31%).

Table 31 *Sensitivity, specificity, PPV and NPV for colorectal cancer development.*

	Sensitivity (%) (95% CI)	**Specificity (%) (95% CI)**	**PPV (%) (95% CI)**	**NPV (%) (95% CI)**
High-risk	2.42% (0.50-6.91)	99.17% (98.08-99.73)	37.50 (12.68-71.25)	83.19 (82.79-83.59)
Moderate & High risk	1.36% (0.50-2.94)	99.31 (97.51- 99.92)	75.00 (37.88-93.66)	39.72 (39.37-40.07)

Table 32 shows the sensitivity, specificity, PPV and NPV of the BSG risk categories for adenomatous polyp development. The combined moderate and high-risk groups have the greater sensitivity (96.18%) but the specificity however is much poorer (12.27%). The high-risk group has a greater sensitivity (18.55%) but lower specificity (93.05%) in comparison to the combined groups.

Table 32 *Sensitivity, specificity, PPV and NPV for adenomatous polyp detection.*

	Sensitivity (%) (95% CI)	**Specificity (%) (95% CI)**	**PPV (%) (95% CI)**	**NPV (%) (95% CI)**
High-risk	18.55% (12.14-26.52)	93.05% (90.72-94.94)	35.38% (25.49-46.71)	84.77% (83.61-85.85)
Moderate & High risk	12.27% (9.36-15.71)	96.18% (93.27-98.08)	83.08% (72.31-90.22)	41.78% (40.77-42.80)

4.3 Kaplan-Meier Analysis

For the KM analysis, the reported *p*-values are the Log-Rank *p*-value.

4.3.1 Cancer

4.3.1.1 Low and moderate-risk group:

Across the entire follow-up time period, two colorectal cancers were developed in the low-risk group and three colorectal cancers in the moderate-risk group. KM survival analysis demonstrated that there is no significant difference in the rate of CRC development between these two groups, *p*=0.681. **Figure 11** shows the KM survival curve.

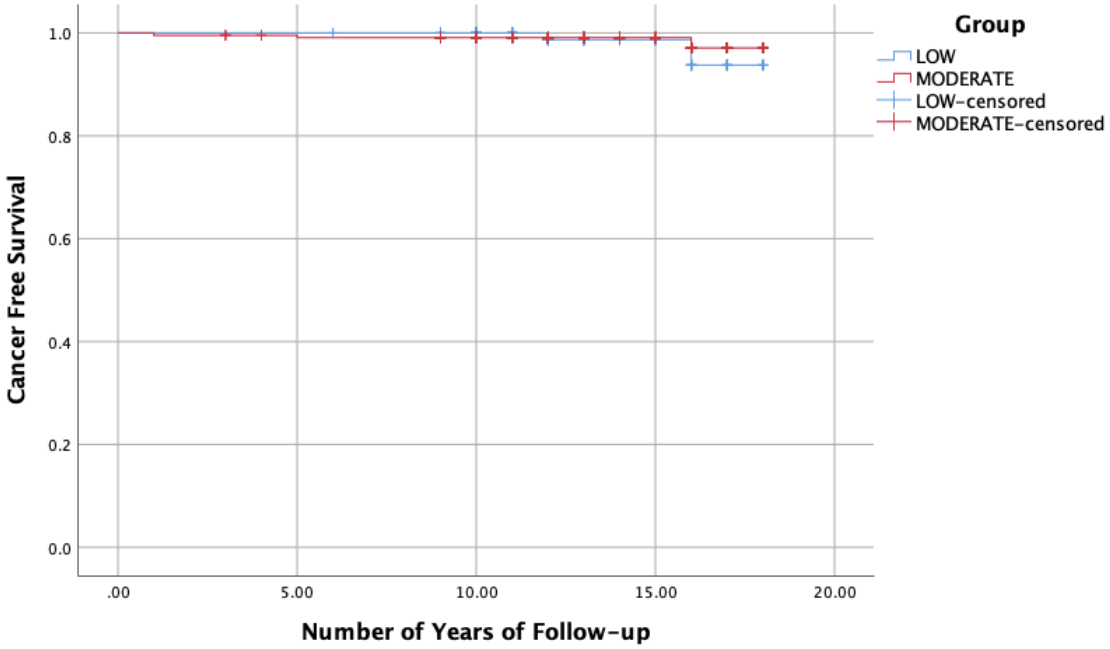


Figure 11 *KM analysis of colorectal cancer in the low and moderate-risk categories across total patient follow-up time.* Log rank p=0.681.

4.3.1.2 Low and high-risk groups:

There were three CRC's in the high-risk group, including two that occurred in the confirmed mutation group, and two CRC's in the low-risk group. There was no significant difference in the KM analysis of cancer free survival when comparing low and high-risk groups, both including and excluding mutation carriers (p=0.540 and p=0.854 respectively). KM survival curves are shown in **Figures 12** and **13**.

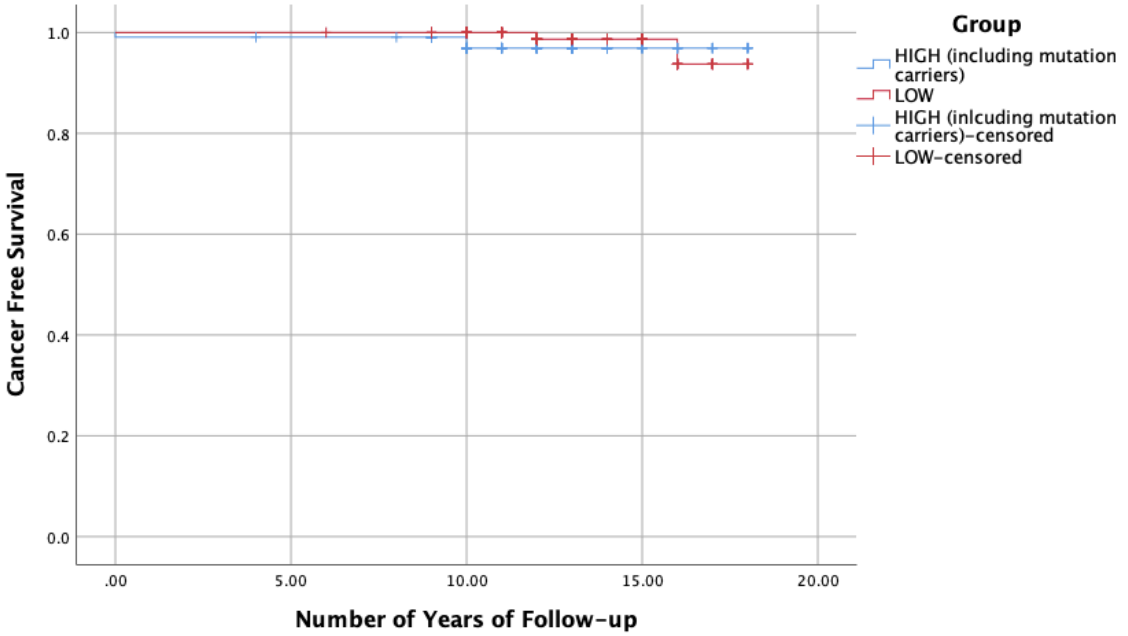


Figure 12 *KM analysis of colorectal cancer in the low and high-risk (including mutation carriers) categories across total patient follow-up time.* Log rank p=540.

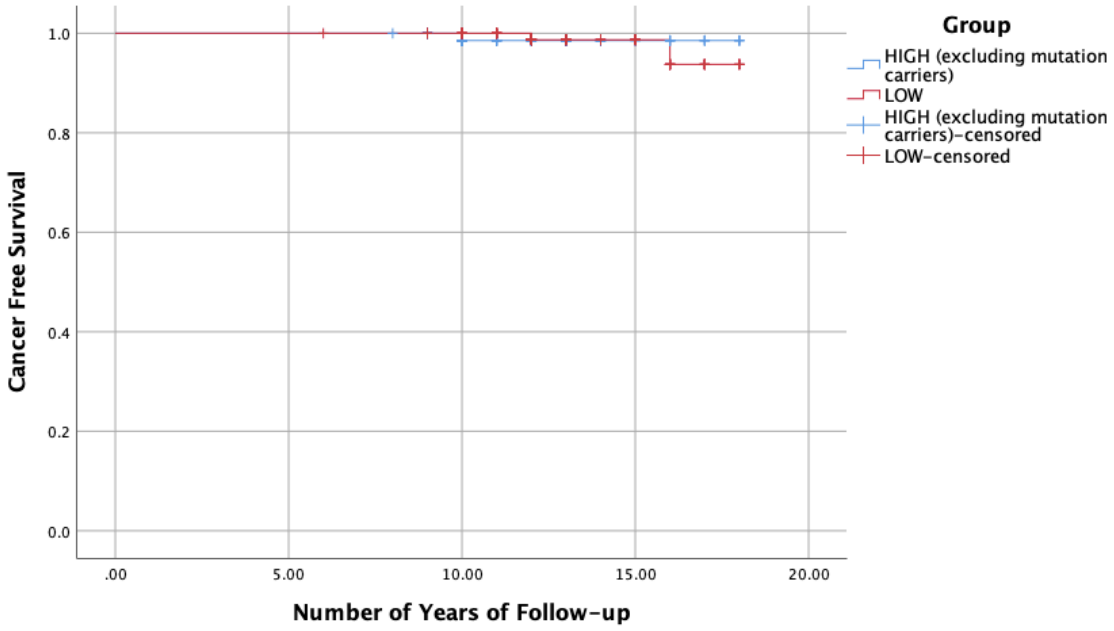


Figure 13 *KM analysis of colorectal cancer in the low and high-risk (excluding mutation carriers) categories across total patient follow-up time.* Log rank p=854.

4.3.1.3 Low and increased risk groups:

Across the follow-up time, there were two cancers in the low-risk group, and six cancers in the increased risk groups (three in each moderate and high-risk groups). When comparing the low and increased risk groups, there was no significant difference in cancer free survival, when including mutation carriers (p=0.995) and when excluding mutation carriers (p=0.697), (*See* **Figures 14** and **15** respectively).

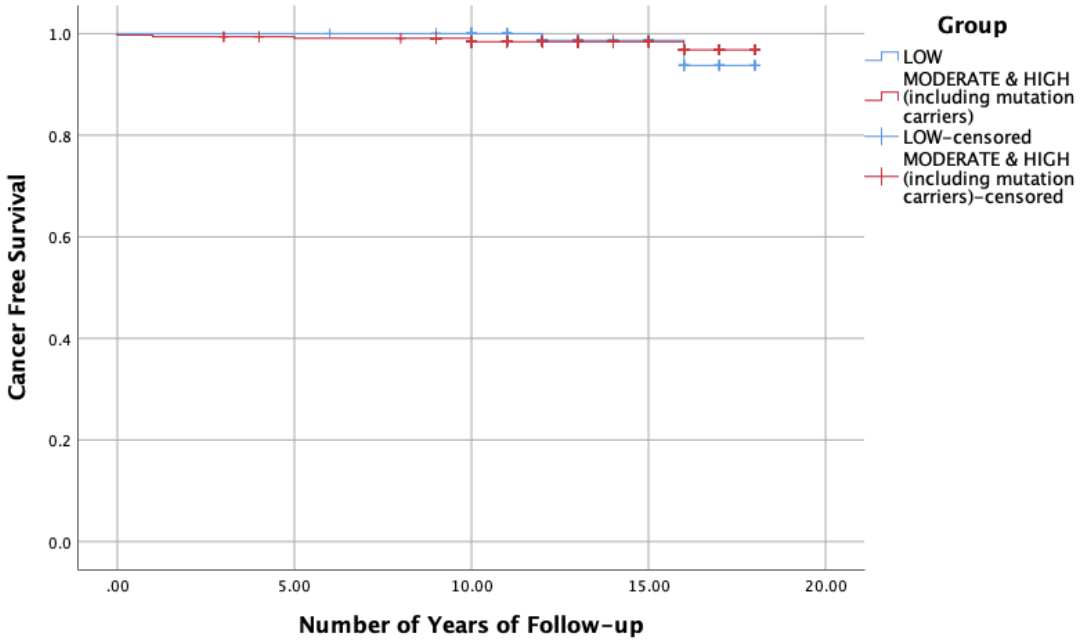


Figure 14 *KM analysis of colorectal cancer in the low and combined moderate and high-risk (including mutation carriers) categories across total patient follow-up time.* Log rank p=0.995.

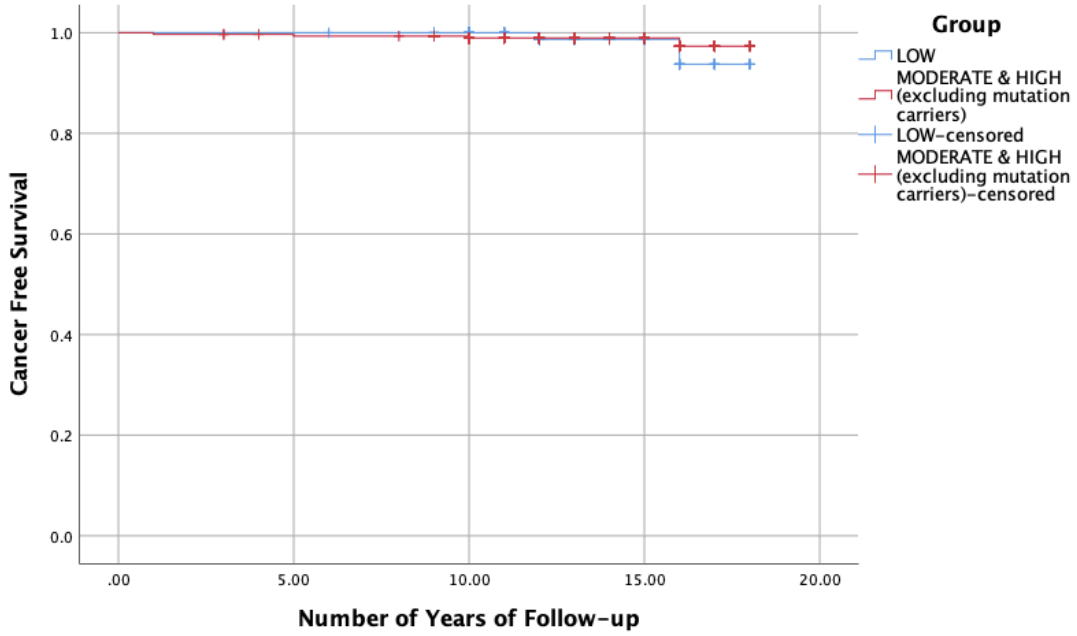


Figure 15 *KM analysis of colorectal cancer in the low and combined moderate and high-risk (excluding mutation carriers) categories across total patient follow-up time.* Log rank p=0.697.

4.3.1.4 Moderate and high-risk groups:

Across all follow-up periods, there were three colorectal cancers events in the moderate-risk group and three colorectal cancer events in the high-risk group, two of which occurred in confirmed mutation carriers. There was no significant difference in CRC development rate between moderate and high-risk groups, including and excluding mutation carriers (p=0.298 and p=0.978 respectively), *See* **Figures 16** and **17**.

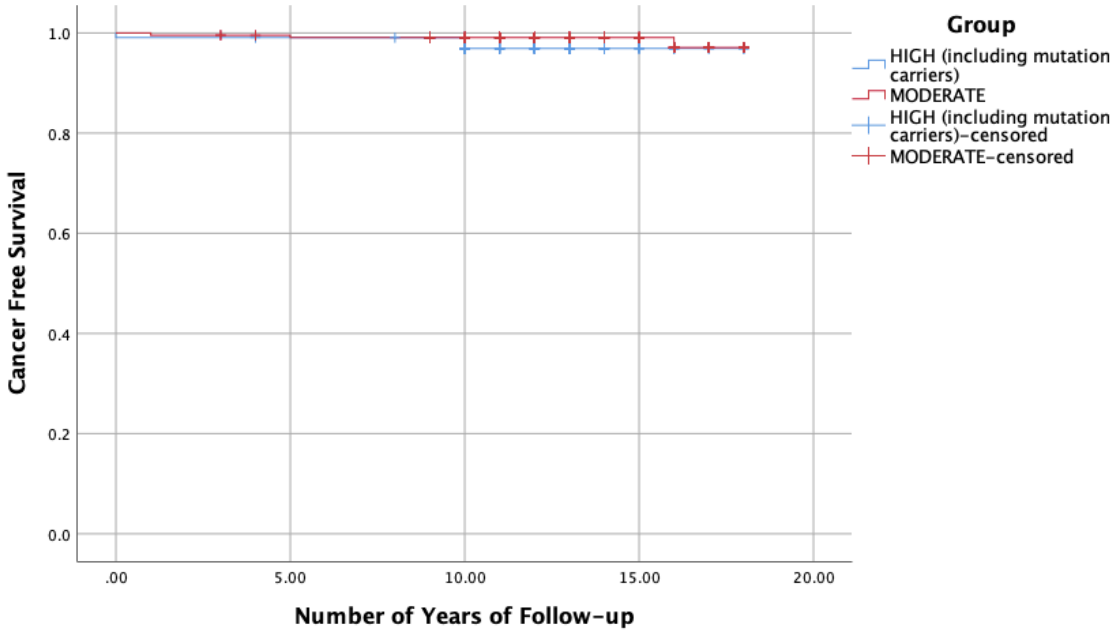


Figure 17 *KM analysis of colorectal cancer in the moderate and high-risk (including mutation carriers) categories across total patient follow-up time.* Log rank p=0.298.

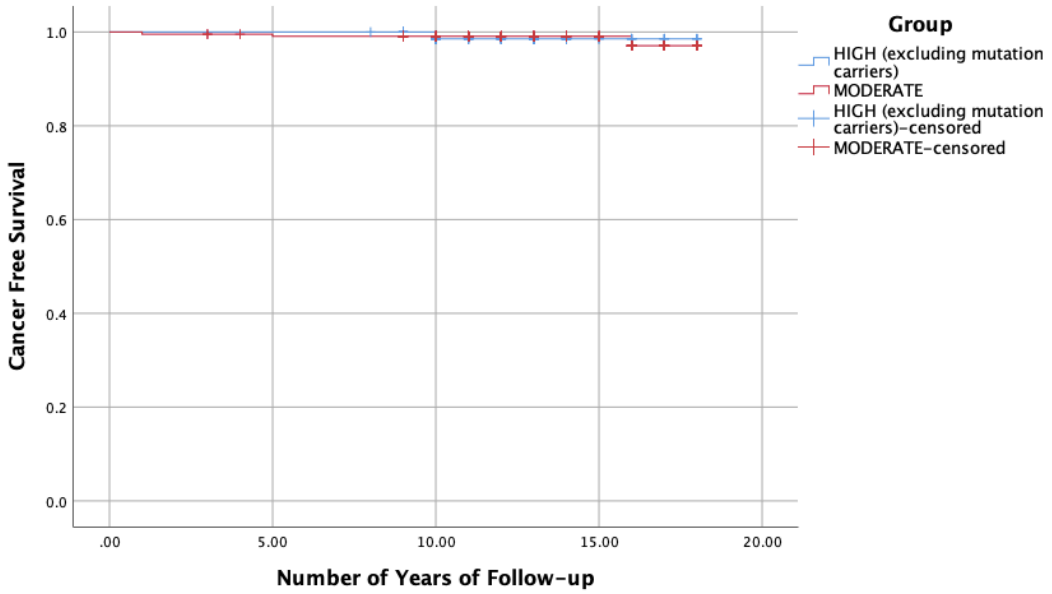


Figure 16 *KM analysis of colorectal cancer in the moderate and high-risk (excluding mutation carriers) categories across total patient follow-up time.* Log rank p=0.978.

4.3.2 Adenomatous Polyps

4.3.2.1 Low and moderate-risk group:

Across all follow-up period, there 11 polyps were detected in the low-risk group and 31 polyps were detected in the moderate risk group. There was no significant difference on KM analysis when comparing polyp detection between these two groups, *p*=0.352. (*See* **Figure 18**).

For the follow-up period for the age-group 0-49 years, there were 8 polyps in the low-risk group and 13 in the moderate-risk group. There was no significant difference in polyp detection between groups, *p*=0.669.

For the follow-up period of ages 50 and above, 13 and 18 polyps were detected in the low and moderate-risk groups respectively. KM analysis was not significant (*p*=0.087).

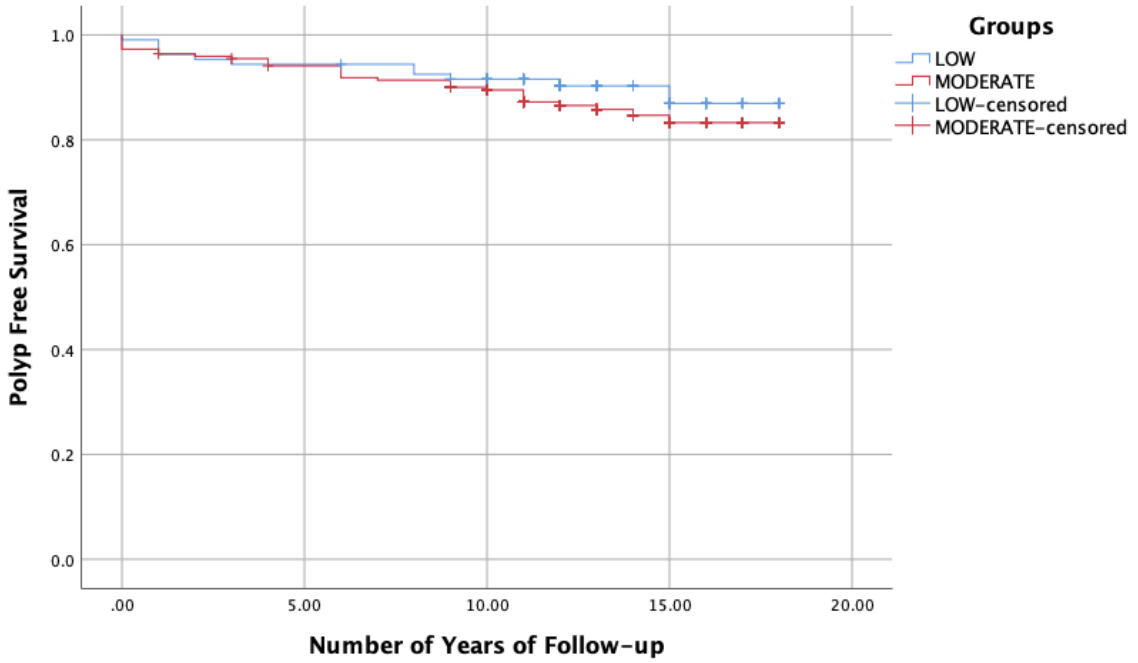


Figure 18 *KM analysis of polyp detection in the low and moderate-risk categories across total patient follow-up time.* Log rank p=0.352.

4.3.2.2 Low and high-risk groups:

There were 23 polyps detected during the follow-up period in the high-risk group, in which 5 occurred in confirmed mutation carriers. There were 11 polyps detected in the low risk group. The likelihood of polyp detection in the high-risk group is significantly higher compared to the low-risk group, both when including and excluding mutation carriers, p=0.017 and p=0.014 respectively, *see* **Figures 19** and **20**.

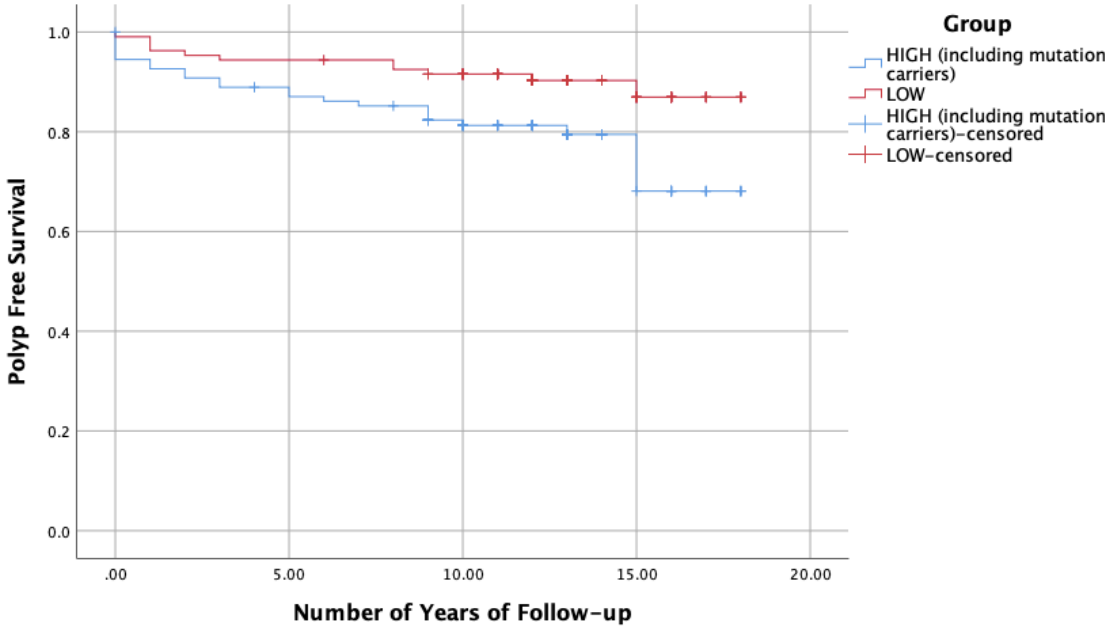


Figure 20 *KM analysis of polyp detection in the low and high-risk (including mutation carriers) categories across total patient follow-up time.* Log rank p=0.017.

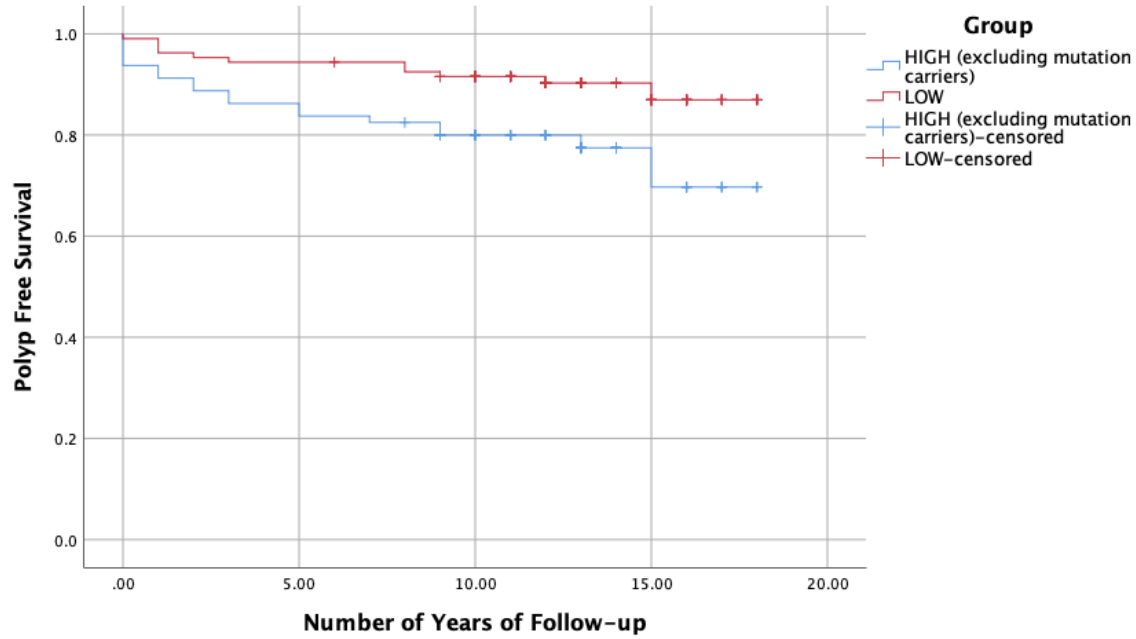


Figure 19 *KM analysis of polyp detection in the low and high-risk (excluding mutation carriers) categories across total patient follow-up time.* Log rank p=0.014.

For ages 0-49, there were 8 polyps in the low-risk group and 18 in the high-risk group, including 5 that were detected in mutation carriers. Polyp detection is significantly higher in the high-risk group compared to the low-risk group for ages 0-49, both when including and excluding mutation carriers (p=0.006 and p=0.008 respectively); *see* **Figures 21** and **22**.

For ages 50 and above, there were 3 polyps detected in the low-risk group and 5 in the high-risk group, with no polyps detected in the mutation carrier group. Polyp detection was not significant between groups, both including and excluding mutation carriers (p=0.584 and p=0.402 respectively).

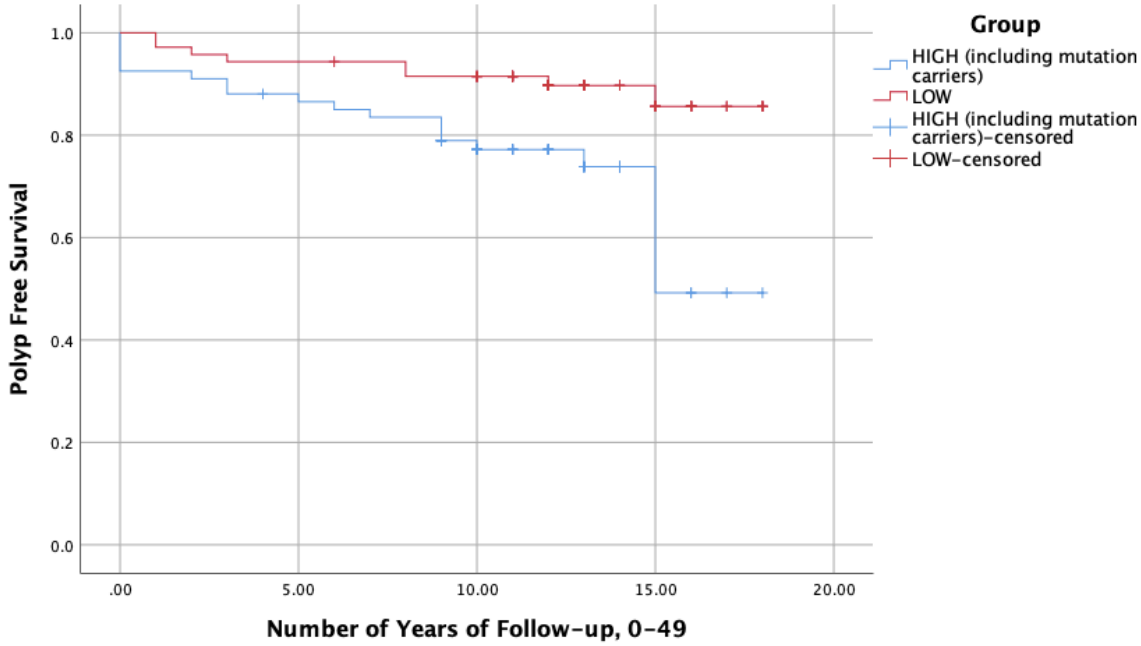


Figure 21 *KM analysis of polyp detection in the low and high-risk (including mutation carriers) categories in ages 0-49.* Log rank p=0.006.

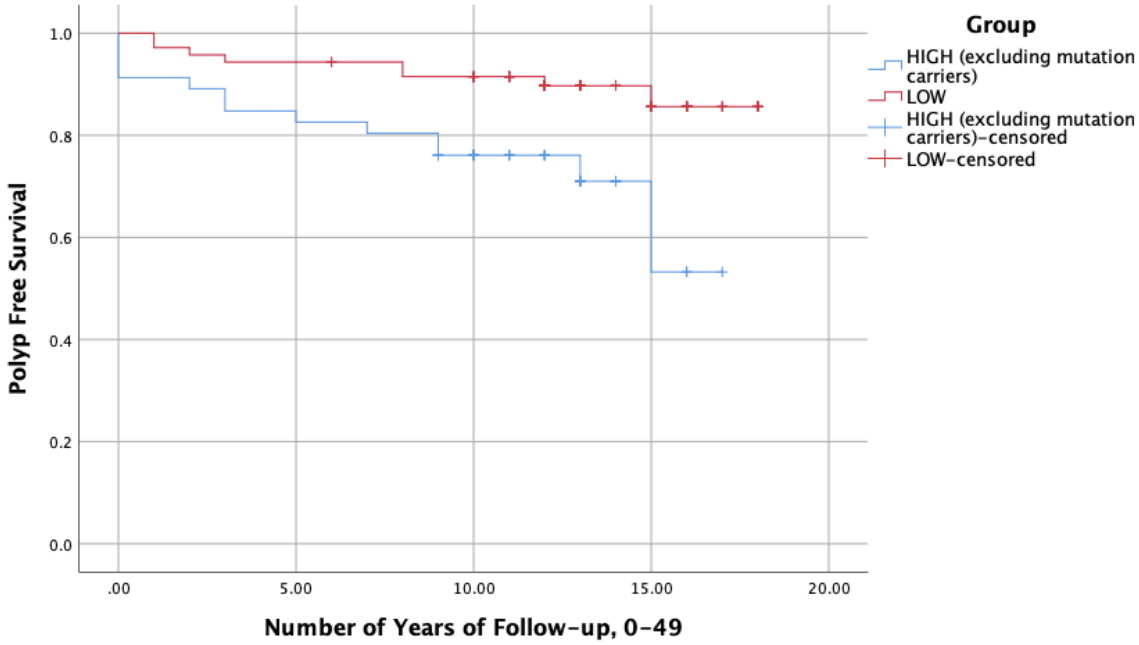


Figure 22 *KM analysis of polyp detection in the low and high-risk (excluding mutation carriers) categories in ages 0-49.* Log rank p=0.008.

4.3.2.3 Low and increased risk groups:

In the moderate and high-risk groups combined, 54 polyps were detected, including 5 in the mutation carrier group. Comparison of polyp detection between the increased risk groups with the low-risk group was not significant, *p*=0.119 (including mutation carriers) and *p*=0.131 (excluding mutation carriers), shown in **Figures 23** and **24**.

Below the age of 50, there were 8 polyps detected in the low-risk group and 31 polyps in the increased risk groups without any statistical significance, including and excluding mutation carriers (*p*=0.384 and *p*=0.523 respectively).

From the age of 50 and above, the increased risk groups had 23 polyps detected, none in the mutation carrier group. Survival analysis of polyp development including and excluding mutation carriers was not statistically significant (*p*=0.155 and *p*=0.123 respectively).

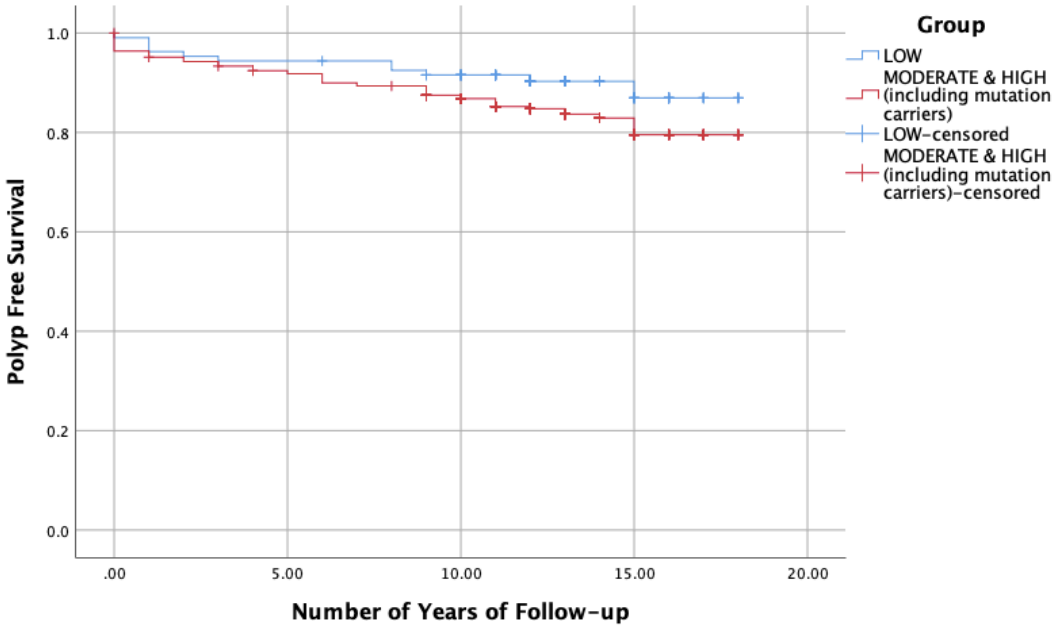


Figure 23 *KM analysis of polyp detection in the low and combined moderate and high-risk (including mutation carriers) categories across total patient follow-up time.* Log rank p=0.119.

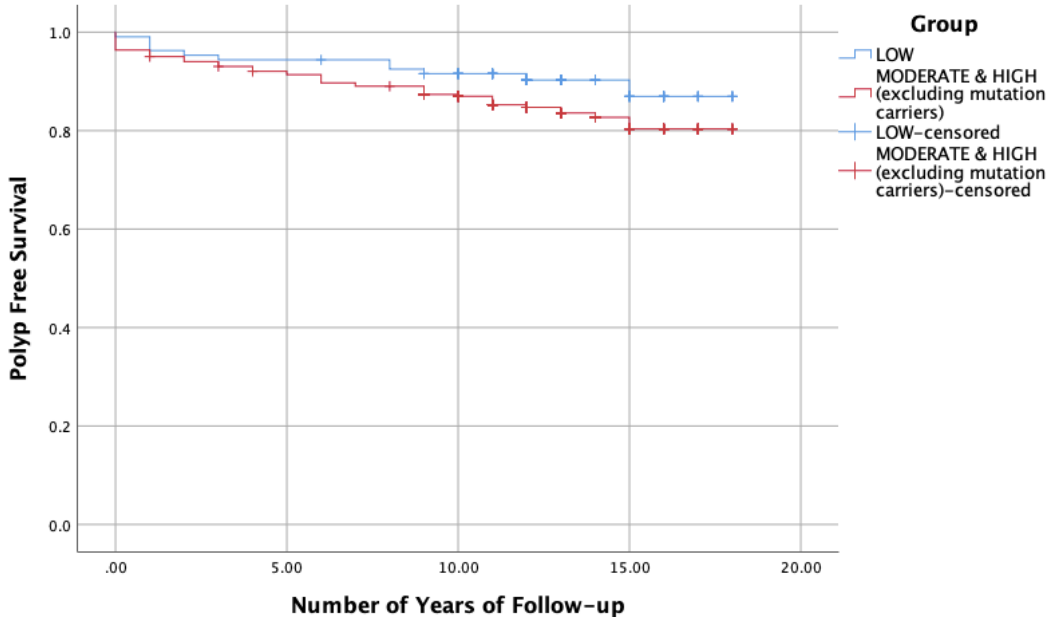


Figure 24 *KM analysis of polyp detection in the low and combined moderate and high-risk (excluding mutation carriers) categories across total patient follow-up time.* Log rank p=0.131.

4.3.2.4 Moderate and high-risk groups:

Across the follow-up period, 31 adenomatous polyps were detected in the moderate-risk category and 23 polyps in the high-risk category, 5 of which occurred in the mutation carrier group. On KM analysis, the rate of polyp detection between moderate and high-risk group was borderline significantly increased in the high-risk group when mutation carriers were included (p=0.053). Rate of polyp detection in high-risk group excluding mutation carriers was significantly higher compared to the moderate-risk group (p=0.042). KM curves shown in **Figures 25** and **26**.

When considering polyp free survival in patients 0-49 years of age, there were 13 polyps in the moderate-risk group and 18 polyps in the high-risk group, including 5 in the mutation carrier group. People in high-risk group have a significantly higher rate of adenomatous polyps compared to the moderate risk group (p<0.001 for both including and excluding mutation carriers); *see* **Figures 27** and **28**.

In patient 50 years of age and above, 18 polyps were detected in the moderate risk group and 5 in the high-risk group, with no polyps in the mutation carrier group. KM analysis showed no significant difference in polyp rate between moderate and high-risk groups for people 50 and above, p=0.208 and p=0.412 for high-risk group including and excluding mutation carriers respectively.

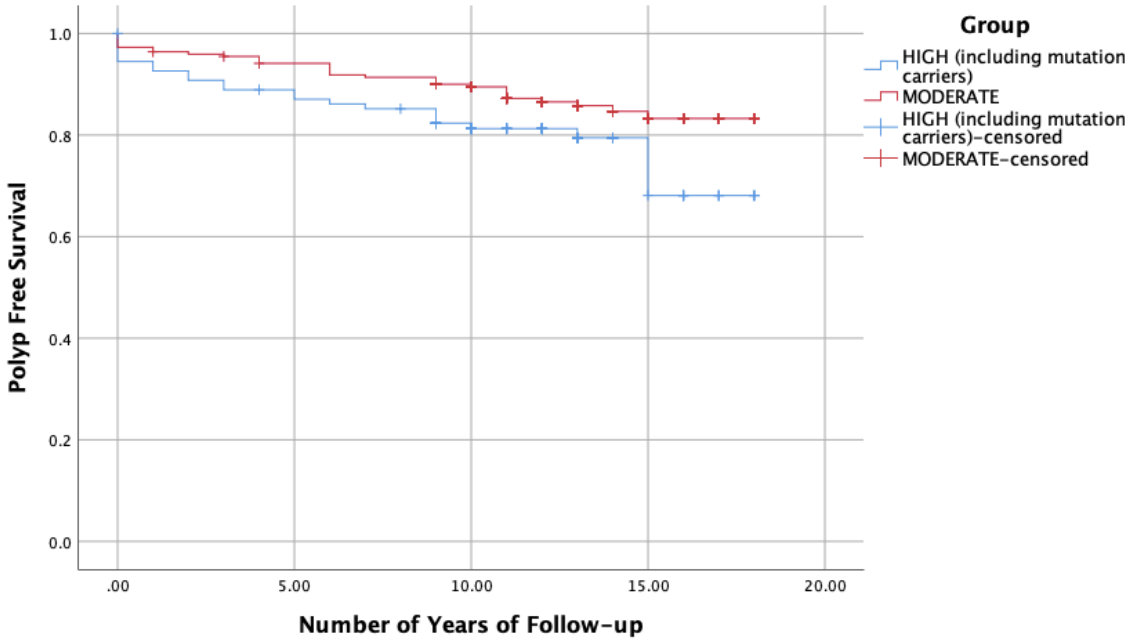


Figure 25 *KM analysis of polyp detection in the moderate and high-risk (including mutation carriers) categories across total patient follow-up time.* Log rank p=0.053.

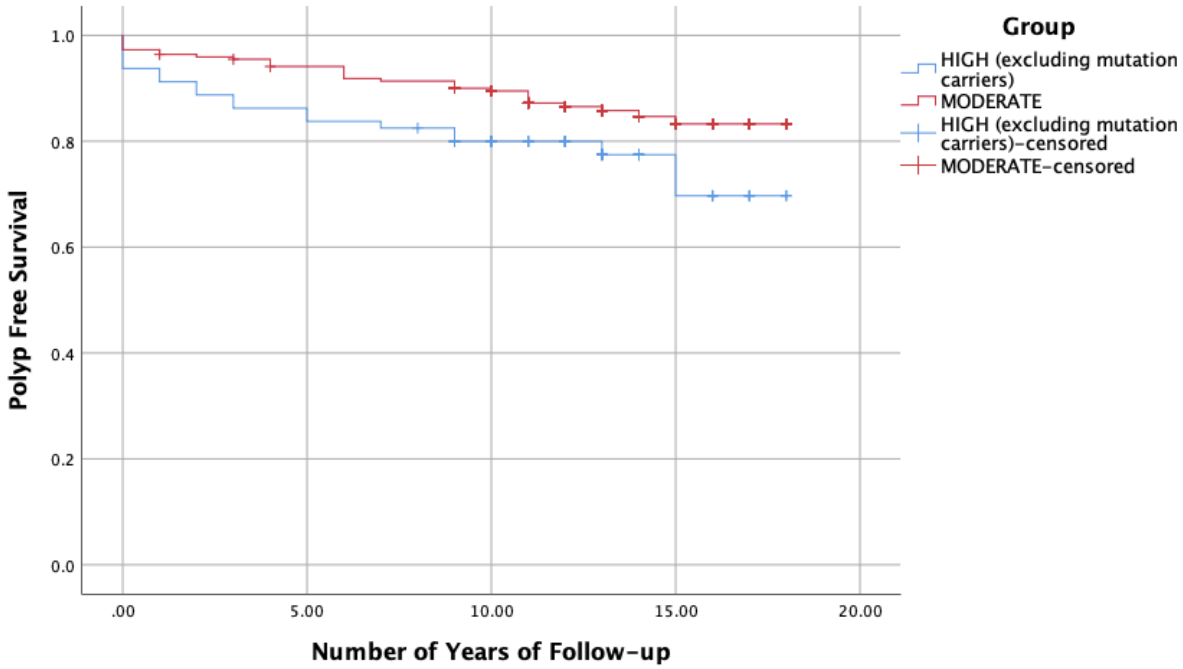


Figure 26 *KM analysis of polyp detection in the moderate and high-risk (excluding mutation carriers) categories across total patient follow-up time.* Log rank p=0.042.

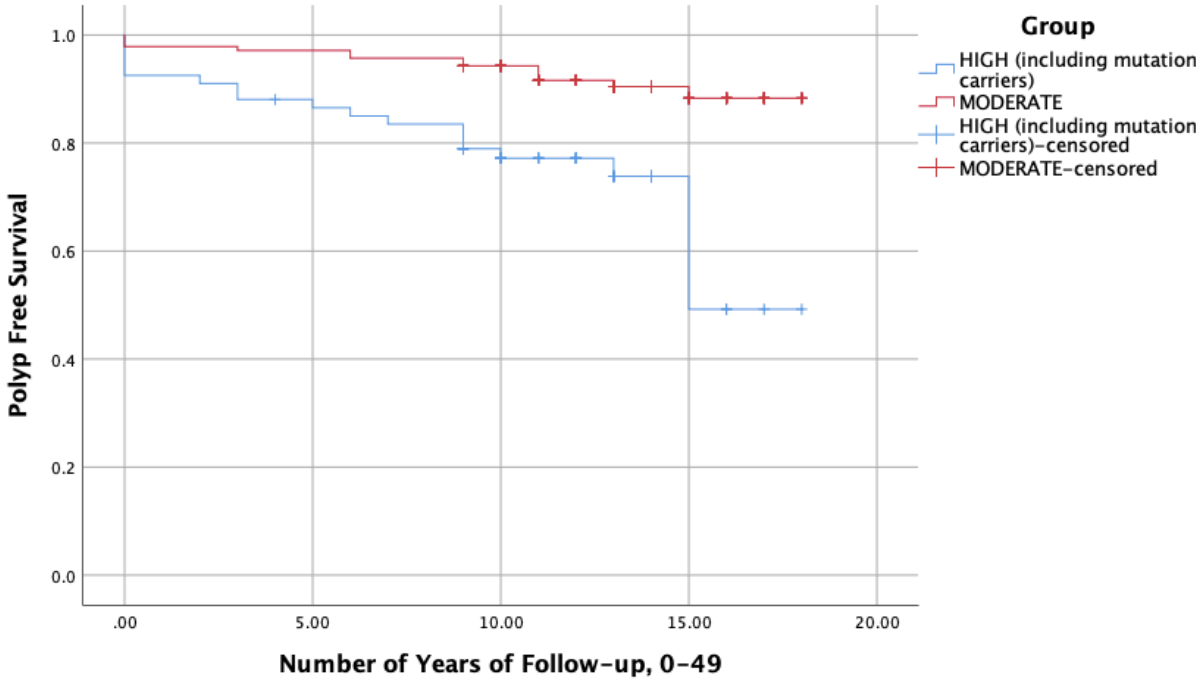


Figure 27 *KM analysis of polyp detection in the moderate and high-risk (including mutation carriers) categories in ages 0-49.* Log rank p= <0.001.

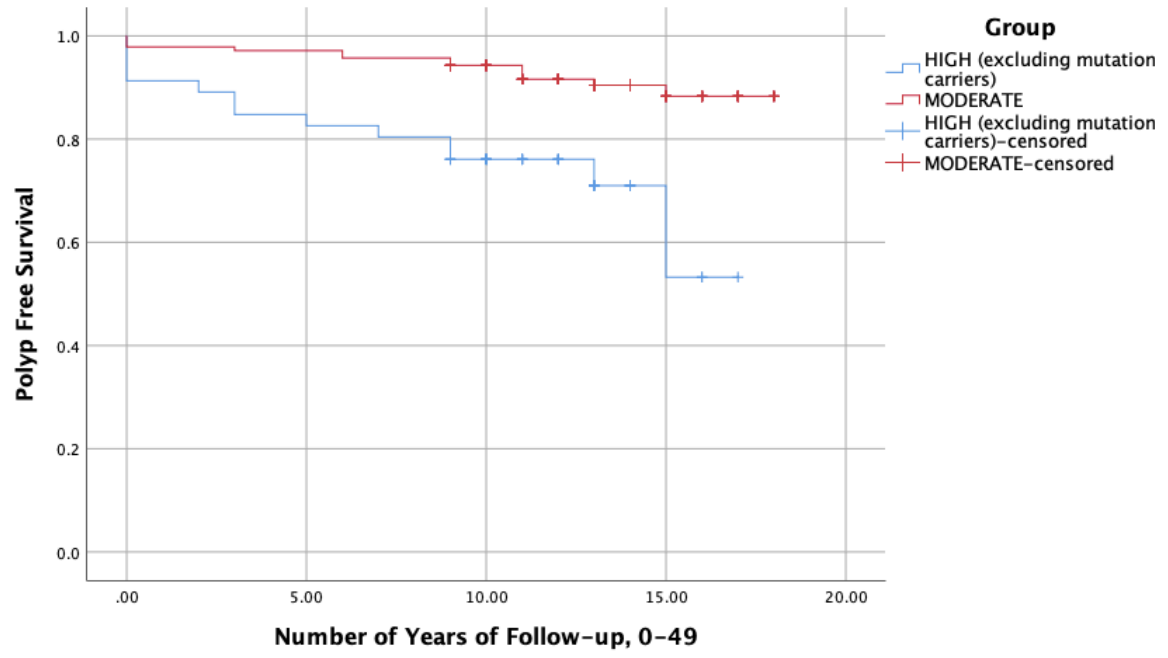


Figure 28 *KM analysis of polyp detection in the moderate and high-risk (excluding mutation carriers) categories in ages 0-49.* Log rank p= <0.001.

4.3.3 KM Summary for Colorectal Cancer:

A summary of the results of KM analysis comparing rates of colorectal cancer diagnoses between different BSG risk categories across all age ranges is shown in **Table 33**. P-values shown are KM Log-Rank p-values.

Table 33 *Summary of survival analysis data for colorectal cancer development.*

BSG Groups Being Compared	**KM Log-Rank (*p-value)***
Low and Moderate	0.681
Low and High (Including mutation carriers)	0.540
Low and High (Excluding mutation carriers)	0.854
Low and Moderate/High (Including mutation carriers)	0.995
Low and Moderate/High (Excluding mutation carriers)	0.697
Moderate and High (Including mutation carriers)	0.298
Moderate and High (Excluding mutation carriers)	0.978

4.3.4 KM Summary for Polyp detection:

Table 34 summarises the results of KM analysis comparing rates of adenomatous polyp detection between different BSG risk categories across all age ranges, 0-49 age group and ≥50 age group. P-values shown are KM Log-Rank p-values. Significant results are highlighted in bold.

Table 34 *Summary of survival analysis data for adenomatous polyp detection.*

BSG Groups Being Compared	**KM Log-Rank (*p-value)***		
	Total Follow-up Time	**0-49**	**≥50**
Low and Moderate	0.352	0.669	0.087
Low and High (Including mutation carriers)	**0.017**	**0.006**	0.584
Low and High (Excluding mutation carriers)	**0.014**	**0.008**	0.402
Low and Moderate/High (Including mutation carriers)	0.119	0.384	0.155
Low and Moderate/High (Excluding mutation carriers)	0.131	0.523	0.123
Moderate and High (Including mutation carriers)	0.053	**<0.001**	0.208
Moderate and High (Excluding mutation carriers)	**0.042**	**<0.001**	0.421

4.4 Summary

Table 35 shows summary information of the frequency and % 10-year absolute risk of CRC, KM analysis and RR for each BSG risk group.

Table 35 *Summary table for colorectal cancer development.*

A)	Number of people	Number of Cancers (% 10-year absolute risk)			KM log rank p-value	Overall RR (95% CI)
		Overall	0-49	>50	Overall	
Low	288	2	1 1.06% (0.468-1.652)	1 2.31% (1.593-3.026)	-	-
Moderate	316	3	0	3 2.92% (2.201-3.640)	0.681	1.367 (0.2301-8.1236)
High (excluding mutation carriers)	93	1	0	1 2.31% (1.372-3.292)	0.854	1.548 (0.1420-16.8830)
Moderate/High (excluding mutation carriers)	-	-	-	-	0.697	-
Mutation carriers only	31	2	1 4.03% (2.714-5.346)	1 10.99% (6.728-10.250)	-	9.290 (1.3557-63.6653)
Total	728	8	2	6		
B)	Number of people	Number of Cancers (% 10-year absolute risk)			KM log rank p-value	Overall RR (95% CI)
		Overall	0-49	>50	Overall	
Low	288	2	1 1.06% (0.468-1.652)	1 2.31% (1.593-3.026)	-	-
Moderate	316	3	0	3 2.92% (2.201-3.640)	0.681	1.367 (0.2301-8.1236)
High (Including Mutation carriers)	124	3	1 1.24%(0.625-1.847)	2 3.82% (2.808-4.829)	0.540	3.484 (0.5894-20.5928)
Moderate/High	-	-	-	-	0.995	1.964 (0.3991-9.6622)
Total	728	8	2	6		

Table 36 shows summary information of the frequency and % 10-year absolute risk of polyps, KM analysis and RR for each BSG risk group.

Table 36 *Summary table for polyp detection.*

A)	Number of people	Number of Polyps (% 10-year absolute risk)			KM log rank p-value			Overall RR (95% CI)
		Overall	0-49	>50	Overall	0-49	≥50	
Low	288	11	8 9.04% (8.177-9.904)	3 7.26% (6.110-8.418)	-	-	-	-
Moderate	316	31	13 7.37% (6.804-7.944)	18 20.18% (19.166-21.193)	0.352	0.669	0.087	2.569(1.3155-5.0148)
High (excluding mutation carriers)	93	18	13 28.32% (26.979-29.661)	5 13.16% (11.500-14.620)	**0.014**	**0.008**	0.402	5.067(2.4845-10.3355)
Moderate/High (excluding mutation carriers)	-	-	-	-	0.131	0.523	0.123	-
Mutation carriers only	31	5	5 21.93% (20.096-23.764)	0	-	-	-	4.223(1.5690-11.3659)
Total	728	65	39	26				
B)	Number of people	Number of Polyps (% 10-year absolute risk)			KM log rank p-value			Overall RR (95% CI)
		Overall	0-49	>50	Overall	0-49	≥50	
Low	288	11	8 9.04% (8.177-9.904)	3 7.26% (6.110-8.418)	-	-	-	-
Moderate	316	31	13 7.37% (6.804-7.944	18 20.18% (19.166-21.193)	0.352	0.669	0.087	2.569 (1.3155-5.0148)
High (Including Mutation carriers)	124	23	18 26.20% (25.143-27.257)	5 10.62% (6.605-14.635)	**0.017**	**0.006**	0.584	4.856 (2.4432-9.6529)
Moderate/High	-	-	-	-	0.119	0.384	0.155	3.213 (1.7094-6.0399)
Total	728	65	39	26				

4.5 Polyposis conditions

4.5.1 Cohort Characteristics

In total there were 52 patients that presented to Clinical Genetics with polyposis conditions. Overall, the age of presentation ranged from 0 years and two months to 69 years. The mean age of presentation was 32.716 (SD ± 18.0630) and the median age was 32 years. All the patients had no history of CRC, and they were excluded if they had undergone any type of colectomy. The total patient years follow-up for CRC was 337. The mean number of patient years of follow-up was 9.629 (SD ± 4.9862), with a median of 11 years.

4.5.2 Family history structures and cancer history

Table 37 summarises the FH data extracted from clinical notes.

Table 37 *FH structure of polyposis cohort.*

	Mean	Range	
		Minimum	Maximum
Sisters (Including half-sisters)	0.88	0	4
Brothers (Including Half-brothers)	1.15	0	6
Total number of siblings	2.03	0	8
Daughters	0.37	0	2
Sons	0.62	0	4
Total number of Children	0.98	0	4
Second Degree relatives	9.42	4	32

Incidence and percentage of cohort with CRC in various family members is described in **Table 38**. The table also reports incidence of gastric and other type of cancers. The most common FDR reported with CRC was the father (17.31%) followed by mother and siblings (11.53% each). Even though polyposis conditions are associated with increased risk of stomach cancer, the only FDR with an incidence of stomach cancer was the father (3.85%).

Table 38 *Cancer incidence in relatives of the polyposis cohort.*

	Incidence in relative amongst cohort (N=52)	
	N	% of cohort
Colorectal cancer incidence		
Mother	6	11.53%
Father	9	17.31%
Sibling(s)	6	11.53%
Child(ren)	1	1.92%
SDR	31	59.62%
Stomach cancer incidence		
Mother	0	-
Father	2	3.85%
Sibling(s)	0	-
Child(ren)	0	-
SDR	8	15.38%
Other cancer incidence		
Mother	7	13.46%
Father	2	3.85%
Sibling(s)	6	11.53%
Child(ren)	0	-
SDR	32	61.54%

Family data regarding age of diagnosis of CRC in the cohort is shown in **Table 39**.

Table 39 *Age of cancer development in relatives of polyposis cohort.*

	Mean (±SD)	Median	Range	
			Minimum	Maximum
Age of total FDRs and SDRs with CRC diagnosis	51.18 (±14.539)	54	26	81
Age of FDRs with CRC diagnosis	44.28(±12.091)	40	26	77
Age of Mother with CRC diagnosis	53.67(±17.461)	49	35	77
Age of Father with CRC diagnosis	38.88(± 9.636)	36	26	62
Age of Sibling(s) with CRC diagnosis	47.33 (±8.260)	50.5	34	55
Age of Child(ren) with CRC diagnosis	41 (±0)	41	41	41
Age of Total SDRs with CRC diagnosis	55.96 (±14.165)	59	26	81

4.5.3 BSG Risk Categories and Mutation Testing

BSG Guidelines assign patients with polyposis conditions into high risk category, but have different screening recommendations according to Family History, Mutation testing and any type of colon surgery. **Table 40** shows the percentage of the cohort that fits into each BSG category. What is not included in the table is FAP patients post colectomy or post procto-colectomy as they were excluded from the study.

Table 40 *Frequency of patients in each risk BSG risk category.*

BSG category	Frequency (N)	% of Cohort
At risk FAP (Member of FAP family with no mutation identified)	7	13.5%
Fulfils clinical FAP criteria	3	5.8%
Proven APC mutation carrier opting for deferred surgery	37	71.2%
MUTYH-associated polyposis (MAP)	1	1.9%
Peutz-Jeghers syndrome (PJS)	2	3.8%
FDR of proven mutation carrier	2	3.8%

From the 52 patients in the polyposis cohort, 1 tested positive for MAP mutations and 2 for PJS and another 37 were APC mutation carriers. From the 37 APC mutation carriers, 19 had mutations that predispose to FAP,14 had mutations that predispose to Attenuated FAP and 2 tested positive for a Variant of Uncertain Significance (VUS). From the cohort, there were 7 patients that had a family member with FAP with no mutation identified and another 3 that themselves fulfilled the clinical FAP criteria with no mutation identified. The complete set of identified mutations and their frequencies is found in **Table 41**.

Table 41 *Confirmed pathogenic mutations for polyposis conditions identified in the cohort.*

Mutations	**Frequency**
Familial Adenomatous Polyposis	
APC exon 13 1660c>t	2
APC c.694C>T,p. (Arg 232*)	2
APC 5757 5772 del 16bp	7
APC c.5979delT	3
APC 5760-5776 del 16bp	1
APC 1753delC	1
APC 288T>A exon 3	14
APC 5461/2delA	1
Apc15.4 exon 3374InsT	1
APC 5461/2delA	1
APC gene deletion	1
APC c 222-2A>g	1
MYH associated polyposis	
MUTYHpGly3965Asp]+[=]	1
Peutz-Jeghers syndrome	
STK11 790delTTGA	1
STK11 c.1529_1533delTCAAA	1
Total	**38**

4.5.4 Number of years of follow-up

From the whole cohort, 35 patients were followed-up for colonoscopic screening. **Table 42** Shows the number of years of follow-up of the polyposis patients until cancer incidence or first adenomatous polyp development, according to age group.

Table 42 *Colorectal cancer and polyp follow-up of patients.*

	0-49 years	≥ 50 years	Total
Cancer Follow-up	295	42	337
Polyp Follow-up	211	29	240

4.5.5 Cancer development and polyp detection in cohort

Overall, no patient developed cancer in this cohort. However, 18 (34.6%) of the cohort had a colectomy either before or after the start of follow-up.

From the 35 patients that were followed up, 18 (51.43 %) of them had adenomatous polyps detected. Those patients first attended clinical genetics with a mean age of 38.6937(SD ± 15.1021), median age 42 and age range of 7-64. The mean time from first appointment to adenoma detection was 3.833 years (SD ±4.5922). The average age of polyp detection in the cohort was 41.325 years (SD ±13.6902), with a median age of 43.

Across the polyposis cohort, no patient developed gastric cancer during the follow-up period. Additionally, only 6 (11.54%) patients have had a gastric polyp identified, in which only two (33.33%) were sessile, the rest (66.67%) were hyperplastic.

4.5.6 Cohort summary

Figure 29 summarises the basic descriptive characteristics of the polyposis cohort.

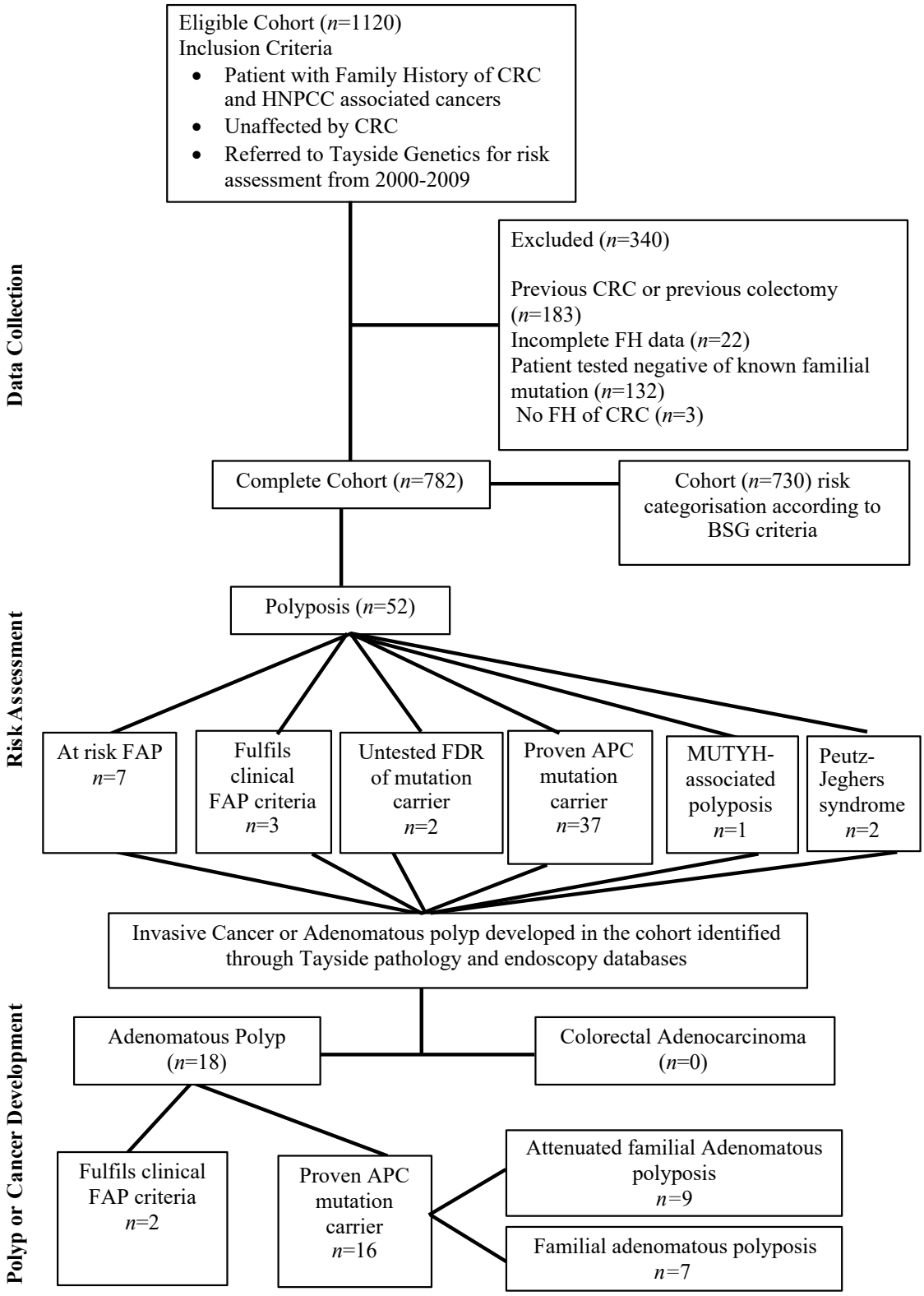


Figure 29 *Summary of selection process and descriptive characteristics of the polyposis cohort.*

Chapter 5: Discussion

5.1 Cohort Characteristics

From the data collection, 782 patients met the inclusion criteria. Of those patients, 728 had a FH of colorectal cancer without an identifiable polyposis syndrome. In total, there were 5561 patient years of follow-up for risk of CRC, with mean number of years of follow-up per person being 12.696 (SD ±2.8092). The patient years of follow-up for polyp detection was 5111, with a mean of 11.669 years (SD±4.0581). Patient years follow-up were less for polyps than cancer, as more polyps were detected during follow-up, thus more patients became censored. The mean age of presentation to clinical genetics was 44.948 years old (SD± 12.2472). When considering family structures and cancer history, a wide variation in presentation was found. Interestingly, the most common FDR presenting with CRC was the mother (34.75%) and with other HNPCC cancers was the father (7.97%). The average age of any FDR being diagnosed with CRC or HNPCC associated cancer is 57.110 (SD±13.0720) years and 54.855 (SD±15.0130) years respectively.

5.2 BSG Risk category

5.2.1 Assigning BSG risk

In total, 288 (39.56%) patients were assigned to a low-risk, 316 (43.41%) patients were assigned to a moderate-risk and 124 (17.03%) patients were assigned to a high-risk category according to BSG criteria, suggesting a reasonable distribution of patients between each group. Since the mean age at presentation of each group was not significantly different, the likelihood of follow-up time at certain ages being significantly different is low. When considering risk categories, the majority (60.44%) of the patients that attended the genetics clinic had an increased risk of developing CRC. In the moderate-risk group, the most common FH presentation was CRC in 1 FDR <50 years of age (13.87%). In the high-risk group, the most common FH presentation was people at risk of HNPCC MMR

carriers that fulfil the Amsterdam Criteria, without an identifiable mutation (11.95%). The family histories used to assign patients their correct risk categories were retrieved from FH questionnaires sent to the patients, pedigrees and clinical letters; where possible, these were confirmed by the relevant cancer registry. Studies have demonstrated that patient-reported cancer family history is accurate when it comes to colorectal cancer risk assessment and that over-reporting is rare. Therefore, although it is important to consider over-reporting biases in different cancers, the risk of over-reporting in the current data and results for colorectal cancer is likely to be low and thus risk assignment is likely to closely reflect the actual risk in this case.(184,185) There may be small discrepancies when a relative had metastatic cancer without primary cancer identification, nevertheless, this did not occur often.

5.2.2 Potential discrepancies in BSG risk criteria

While assigning patients to the BSG risk categories, some potential pitfalls were identified. Firstly, the guidelines were interpreted in the literal sense for the purposes of the study as the aim was to assess the utility of the current risk stratification guidelines. Therefore, for example, the criteria for moderate-low risk group include CRC in 1 FDR <50 years. Thus, a patient with a FDR with CRC diagnosed at 51 may be placed into the low-risk category even though their risk may be the same as the person diagnosed at 49. Furthermore, the criteria only included First- and Second-degree relatives in a first-degree kinship. It can be argued that a relative more distant than second-degree does not share enough genetic material with the patient to be relevant to a polygenic mode of inheritance. However, since the penetrance of CRC and other HNPCC associated cancers is not 100%, it may be that the patient's aunt is yet to develop the condition, in which a higher risk may be re-assigned at a later stage.

Furthermore, even though both age and number of relatives are included in the risk stratification, it can be argued that the BSG guidelines use age as the key determinant when assigning a patient to a risk group, as family structure is not considered. This is because the guidelines fail to address the potential impact of

the size of the family in risk stratification. For example, a person whose mother is affected with CRC and also has two siblings, both affected with CRC may be at a greater risk from a relative whose mother is affected with CRC and has six siblings, in which only two of them are affected with CRC. The model also fails to take into account the gender of relatives. BSG criteria state that the lifetime risk of CRC is higher in males (1 in 5) rather than in females (1 in 13), which may result in misplacing patients in a higher or lower risk category, depending on their relatives' gender.(37)

Similarly, there may be cases where there is one CRC and two other HNPCC associated cancers, none less than the age of 50, in which the patient does not meet Amsterdam and therefore high-risk criteria. The same patient will not even meet the moderate-high risk criteria, as there is no mention of other HNPCC cancers, except for CRC. In cases such as the above, the judgement of the clinician and genetic counsellor will determine the risk category in which the patient will be assigned to. These exceptions allow for subjective interpretation of the guidance amongst different centres, which may limit the reliability of risk categorization and have a downstream effect on the follow-up screening recommendations. A universal guidance that takes into account some of these discrepancies would have been desirable to avoid this element of uncertainty.

5.3 Mean age of CRC development and polyp detection amongst different risk groups

5.3.1 Mean age of cancer development is not significantly different between groups

The mean age of the patients at CRC diagnosis for each risk group is shown in **Table 23, Section 4.2.1**. One-way ANOVA demonstrated no significant difference between age of presentation of CRC across the risk categories. Even though the high-risk group would have been expected to have a lower age of CRC diagnosis, increased screening in the high-risk group potentially results in successful prevention of cancers at a younger age.

5.3.2 Mean age of Polyp detection is significantly different between groups

The age of adenomatous polyp detection is significantly lower in the high-risk group relative to the moderate-risk group. This is consistent with previous studies, which have found that people with familial CRC are at risk of adenoma formation at an earlier age.(172,186) This may also explain the above findings of no difference in age presentation of CRC's between groups, as if polyps are detected early through colonoscopic screening in the high-risk patients, they can be adequately excised and followed up, reducing CRC development. Therefore, this result supports the recommendation of colonoscopy being performed from an earlier age in high-risk groups.

5.4 There is greater risk of CRC development in the mutation group

As shown in **Table 27, Section 4.2.5**, eight cancers developed overall, with two cancers in the low-risk group, three in the moderate-risk group and three in the high-risk group. Two of the cancers occurred in patients less than 50 years of age (one in the low-risk and one in the mutation group), whereas the rest occurred in patients ≥50 years old.

In this cohort, the % 10-year absolute risk (AR) from age 0-49 for each category was 1.06% for the low-risk group and 1.236% for the high-risk group (including the mutation carriers). The mutation carrier group on its own had a 10-year % AR of 4.03%. In the age group ≥50, the % 10-year AR for development of CRC for the low, moderate and high-risk was 2.31%, 2.92% and 3.82% respectively. The mutation carrier group on its own had a % 10-year AR of 10.99%.

Thus, the mutation group has a higher risk for CRC development compared to the rest of the groups. Furthermore, the RR and OR for CRC development was significantly higher in the mutation carrier group. This is consistent with other studies that showed that people with Lynch syndrome have a higher risk of CRC development.(108,109)

5.5 There is no difference in risk of CRC development between low and high-risk groups

In ages 0-49, there was no CRC development in the high-risk group. For people aged ≥50, the % 10-year AR for development of CRC for the low-risk group and for the high-risk group (excluding mutation carriers), was 2.31% for both. Additionally, the RR and OR for CRC development was not significant in either of these categories.

From available data, there is evidence that CRC risk is greater in the high-risk group compared to the general population. The fact that there is no difference in CRC risk between these groups across the two age categories suggests that the recommended screening surveillance is effective in reducing the risk of CRC in high-risk groups.

5.6 There is a greater risk of polyp detection in high-risk groups compared to the low-risk group

Overall there were 65 polyps detected across all age groups, 11 in the low, 31 in the moderate and 18 in the high-risk groups. In the 0-49 age category, 39 polyps were detected and in the ≥50 age category, 26 polyps were detected.

In the present study, the % 10-year absolute risk of polyp detection for the low, moderate and high-risk groups were 9.04%, 7.37% and 26.20% respectively for the age group 0-49. The latter value included mutation carriers, whereas the risk rises to 28.32% when the mutation carriers are excluded. Overall, the greatest risk for polyp detection was in the high-risk group, which is consistent with other studies.(172,186) For ages ≥50, the risk was 7.26% for the low, 20.18% for the moderate and 10.62% for the high-risk group (and 13.62% when excluding mutation carriers). It is also worth noting that no mutation carriers developed polyps in this age group, which may be explained in a number of ways. Firstly, mutation carriers may be predisposed to develop polyps at an earlier age, as 5 polyps were detected in the 0-49 age group. Secondly, the increased

colonoscopic surveillance before the age of 50 may account for the earlier polyp detection and excision.

Moreover, there was an increasing trend for RR and OR of polyp detection across BSG categories. The risk for polyp detection was significantly higher for the high-risk group excluding mutation carriers. The results demonstrate that there is an increased risk of polyps in the increased-risk groups compared to the low-risk group across all age categories, which is consistent with the findings of other studies.(186) The risk also seems to be greater in the 0-49 age category.

5.7 There is no significant difference in the rate of CRC development between BSG risk groups

To determine the difference in follow-up time for CRC diagnosis between groups, KM analysis was used. The number of CRC cases in the low and medium-risk groups were two and three respectively. In the high-risk group there were three CRC diagnoses, two of them which occurred in mutation carriers. There was no significant difference in CRC development rates when comparing low and medium, low and high as well as moderate and high-risk groups. Furthermore, no significant difference was detected when comparing low and increased risk groups (i.e. moderate and high-risk groups combined). Separate analysis including and excluding mutation carriers from the high-risk group yielded no significant differences in the rate of CRC development when compared to the moderate-risk group.

From the above findings, it can be argued that the rate of CRC development in the high-risk group is no different compared to the low-risk group, thus, two-yearly colonoscopic screening is not required. On the other hand, it can be argued that the fact that there is no difference in the rate of CRC is due to the increased colonoscopic screening. In order for these hypotheses to be tested, a randomised controlled trial (RCT) should be conducted where half of the high-risk group is exposed to regular colonoscopic screening, in comparison to the other half which does not. However, this would not be realistic due to the ethical implications of

not screening the high-risk group which are known to have a higher risk of cancer development.

5.8 Polyp detection rate is significantly higher in the high-risk group compared to the low and moderate-risk groups, especially in the 0-49 age category

When comparing low and high-risk groups, the rate of polyp detection was significantly greater in the high-risk group across all follow-up time. This was true both when including and excluding mutation carriers from the high-risk category. When age-dependent analysis was performed, it showed that there was a significantly higher risk of polyp detection in the high-risk group, only in the 0-49 age group. The rate of polyp development was significantly higher in the high-risk group compared to the moderate-risk group, both when including and excluding mutation carriers across all follow-up period. When age-group specific analysis was performed, there was a significantly higher rate in polyp detection in the high-risk groups in the 0-49 age group, when excluding mutation carriers.

Even though the above findings showed that there is no difference in CRC risk and rate between low and high-risk groups, polyp risk and rate are greater in the high-risk group. Thus, it can be argued that colonoscopic surveillance is necessary to identify adenomatous polyps at an early stage and prevent adenoma-carcinoma development.

5.9 Medium-risk patients do not require increased surveillance before the age of 50

Moderate-risk group had no CRC development and the lowest 10-year % AR of polyp development compared to the rest of the groups. RR and OR showed that CRC development in the moderate risk group is not significantly higher compared to the low-risk group and that risk of polyp detection is less than the high-risk group. KM analysis showed that there is no significant difference in CRC development or polyp detection rate between the low and the moderate-risk

groups. Additionally, there was a significantly greater likelihood of polyp development in the higher-risk group compared to the moderate-risk group in 0-49 age category.

The above findings indicate that colonoscopic screening in the moderate-risk group before the age of 50 might not be required. This is consistent with the BSG guidelines which do not recommend any screening prior to this age for the moderate-risk group. In order to be able to draw further conclusions regarding the recommended screening frequency, analysis of the moderate-risk group should be separated into moderate-low and moderate-high risks, due to the fact that these two sub-categories have different screening recommendations.

5.10 Further work is required to demonstrate the benefit of screening in high-risk patients aged 50 years or more

From the present cohort, no difference was found in CRC rates between moderate and high-risk groups, both including and excluding the mutation carriers. This is conflicting with the results on polyp rates, which demonstrates that the high-risk group has a significantly higher rate of polyp detection compared to the low and moderate-risk groups, especially in the 0-49 age group. However, these data may not indicate a true connection between high-risk patients and polyp development, but rather a finding caused by the frequent screening of the high-risk group compared to the other two groups.

After the age of 50, survival analysis showed no difference in the rate of polyp detection in the high-risk group when compared to the low and moderate-risk groups. Furthermore, the RR and OR for CRC development in the high-risk group excluding the mutation carriers were non-significant when compared to the low-risk groups and the % 10-year risk of CRC was exactly the same as the low-risk group.

From the above findings, it may be reasonable to decrease the frequency of screening in patients at high-risk (excluding mutation carriers) after the age of 50.

Nevertheless, this should be firstly confirmed by having a larger patient cohort, to allow for smaller age-group analysis and thus give more information regarding CRC risk of high-risk patients above the age of 50.

5.11 Current family history guidelines are not good predictors of cancer risk

No conclusive predictors could be identified from the data that would increase cancer risk. This was true when including and excluding mutation carriers. Moreover, analysis of the high-risk and moderate and high-risk groups combined showed that the categories have very poor sensitivity, even though specificity is good. This means that theoretically, people might have been placed in the low-risk category due to FH, even though they belonged in a higher risk group.

As seen from the results, higher risk groups have an increased risk and rate of adenomatous polyp detection. Adenomas may be detected and excised, preventing cancer development. Thus, FH of adenomas could be potentially incorporated into the guidelines. For a more accurate risk stratification, FH as well as environmental factors that are known to increase CRC risk should be used in a polygenic risk assessment model.

5.12 Screening recommendations for polyposis patients is effective

In this study, only 52 patients were identified with polyposis conditions. 18 of those patients had a colectomy either before or after the start of follow-up. From the patients that were followed-up, nobody developed colorectal cancer, despite the increased risk in polyposis mutation carriers. Even though no risk or survival tests were carried out, the BSG guidelines seem to be give effective recommendations in regard to this risk group and achieving the reduction in cancer.

A prospective study with similar methodology should be conducted for polyposis patients including those with partial colectomies to identify risk of polyps and colorectal cancer and further evaluate screening recommendations.

5.13 Strengths and limitations of the study

5.13.1 Strengths

To the authors' knowledge, this is the first study to attempt to assess the recommendations of the BSG guidelines. The methodology presented in the study allows to assess the effectiveness of the guidelines considering a comprehensive list of variables such as CRC and polyps, and FH taking into account CRC and other HNPCC cancers. Unfortunately, it was not feasible to have a larger cohort due to time constraints, but the design of the study is relatively simple and adaptable, so this can be replicated in other centres with larger cohorts. The method can also be replicated to expand the analysis on polyposis, assessing the effectiveness of the guidelines if there is a larger cohort.

Another strength of the study is the FH information gathered for each individual. The database used provides a detailed structure of FH for a very large cohort of patients, which can be used in the future for continuation of this work or other types of research. The fact that the age of presentation to the clinic was not significantly different across the groups means that the results were not biased in terms of age. The medium-risk group had the most patients, followed closely by the low-risk group. The high-risk group had around half the patients compared to the other two groups and the mutation carriers consisted only a small proportion of that group (25%) which would be the trend expected in a population.

Furthermore, the study is adequately powered to detect differences of colorectal cancer risk between low and moderate and between low and high-risk groups, as indicated by the power analyses. There were also 5561 years of follow-up which are considered a sizable amount for this patient cohort.

5.13.2 Limitations

Firstly, this cohort study was performed retrospectively. People eligible for this study had no personal history of CRC themselves and the FH information used was collected from pedigrees and clinic communications. The fact that the patients were sent out family history questionnaires minimises the risk of recall

bias during the appointment, as they had time to investigate and ask relatives. However, HNPCC is associated with a number of different types of cancer and sometimes it can be difficult to identify all of them during family history. Nevertheless, when permission was granted, cancer registries were used to identify the exact type of cancer and age of relatives, adding an extra layer of accuracy in the reported histories. Despite these limitations, family-history questionnaires are thought to be accurate in collecting information for CRC.(184,185)

Due to the fact that the FH were taken at initial assessment, changes in FH after further mutation testing or further cancer development were not taken into account. Thus, the risk assignment of some people might have changed over the years. This limitation could have been avoided in a prospective study, recruiting patients and following them up over many years. Nonetheless, the aim of the study was to assess the effectiveness of the guidelines at assigning risk of the patients, and thus initial risk assessment is important.

Another limitation of the study was the fact that some information that would have been useful for the study was not present in the clinical notes. Patients with more than 50% of data missing were excluded from the study, decreasing the sample size. In a prospective study design, this could have been avoided by ensuring specific FH protocols and improved history taking. Due to time limitations, prospective data collection to address the study aims was not feasible. As the BSG guidelines recommend excision of all adenomatous polyps, regardless of size, it was not possible to assess the adenoma-carcinoma progression. However, this link is well established and just separate analysis of adenoma detection in the cohort might be deemed sufficient.(172)

Due to the low number of cancers in each group, it was not feasible to divide the population into smaller age groups for risk analysis. Colorectal cancer rates were analysed using the categories of 0-49 and ≥50 years of age, which are relatively broad. These categories seemed appropriate as early colorectal cancer is generally defined as being developed in ages <50. Dividing the ≥50 category into 50-59 and ≥60 would have been more desirable to identify whether there is a

difference in smaller age groups, had there been more cancer incidences in the cohort. If the sample size and follow-up years were larger along with the cancer and polyp incidence, it would have been viable to also calculate the % 5-year absolute risk.

Another important limitation of the study was the fact that it was underpowered when comparing moderate and high-risk groups, thus the results must be interpreted with caution. Nevertheless, the study was still adequately powered to identify risks between low and moderate and between low and high-risk groups. As a pilot study, it still demonstrates the methodology that can be applied to a population with an adequate power to produce more confident results.

5.14 Conclusions

The results suggest that the mutation carrier group have a higher risk of CRC development, thus should be screened regularly, as per the current BSG guidelines. The risk and rate of CRC development does not differ between low and high-risk groups, indicating that the increased screening in the high-risk patients is effective. However, it may be reasonable to reduce screening in the high-risk group (excluding mutation carriers) above the age of 50.

Furthermore, the results re-affirm that screening for the moderate risk group below the age of 50 may be unnecessary, and there is some evidence to suggest that the high-risk group may be eligible for less frequent screening. There is no evidence to suggest less frequent screening in the mutation carrier group.

This study demonstrates a feasible methodology that could be expanded to other genetic centres in Scotland in order to generate a large sample size for addressing the research questions. If replicated, the results presented could have implications for screening recommendations in terms of how at-risk people are identified, and therefore early detection and prevention strategies.

5.15 Future work

5.15.1 Increased sample size

In order to be able to answer questions regarding the effectiveness of risk assignment and screening recommendations with certainty, this study needs to be adequately powered and performed in a prospective fashion. Retrospective power calculation showed that a sample size of 4000 people would be adequate, ideally spread evenly between groups. Since there is a difference in lifetime risk of CRC amongst genders, it would be preferable for the power calculations to be tailored to males and females. In Tayside, only 1120 patients presented to clinical genetics enquiring about their risk of colorectal cancer in a ten-year period. Thus, it is likely that collection of these data from larger centres across the whole of Scotland would achieve a substantial cohort size. A multi-centre approach would require communications with specialists to interpret the guidance in a standardised way, however, developing these guidelines is beyond the scope of this study. The information gathering stage of this study took approximately 8 months part-time, including having to extract data from different databases. Depending on accessibility of local databases, it would be possible to collect data from a larger cohort in a reasonable timeframe. Ideally, a cohort study such as this should be conducted prospectively, however this would require several years of follow-up. Furthermore, recording of FH should be complete and double-checked from available registries if possible, to create a reliable database identifying families at an increased risk.

5.15.2 Survival and cost-benefit analysis

The main outcome of the screening recommendations is to reduce mortality of people with increased risk of colorectal cancer, rather than just identify them. This cohort only identified eight patients with CRC, which are not enough to analyse survival.

The study should also be extended to include a cost-benefit analysis of the screening programme, to identify the cost to run the colonoscopies so frequently and the overall benefit to the patients. This can also be compared with cost-benefit analysis from the Scottish Bowel screening programme. If the cancer free survival of the increased risk patients is significant, this would outweigh the cost.

5.15.3 Improving sensitivity of the guidance

The sensitivity of the guidance in this cohort was poor, thus improving this would also be beneficial. In order to avoid false negatives and improve the sensitivity of the study, it may be worthwhile looking at other family members and exploring history in more detail. Sensitivity can also be improved by including data on adenomatous polyps of relatives.

References

1. Gco.iarc.fr. (2018). *Cancer today*. [online] Available at: http://gco.iarc.fr/today/fact-sheets-cancers [Accessed 9 Jun. 2020].
2. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D., Forman, D. and Bray, F. (2014). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, 136(5), pp.E359-E386.
3. Cancer Research UK. (2019). *Bowel cancer incidence statistics*. [online] Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence [Accessed 2 Jan. 2020].
4. Isdscotland.org. (2019). *Cancer | Publications | Health Topics | ISD Scotland*. [online] Available at: http://www.isdscotland.org/Health%2DTopics/Cancer/Publications/ [Accessed 2 Jan. 2020].
5. Ons.gov.uk. (2018). *Cancer registration statistics, England Statistical bulletins - Office for National Statistics*. [online] Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/previousReleases [Accessed 2 Jan. 2020].
6. Cancer Research UK. (2019). *Bowel cancer mortality statistics*. [online] Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/mortality#heading-Zero [Accessed 2 Jan. 2020].
7. Isdscotland.org. (2018). *Cancer | Cancer Statistics | Colorectal Cancer | Health Topics | ISD Scotland*. [online] Available at: https://www.isdscotland.org/Health-Topics/Cancer/Cancer-Statistics/Colorectal/#colorectal [Accessed 2 Jan. 2020].
8. Mandel J, Church T, Bond J, Ederer F, Geisser M, Mongin S et al. The Effect of Fecal Occult-Blood Screening on the Incidence of Colorectal Cancer. New England Journal of Medicine. 2000;343(22):1603-1607.
9. Hardcastle J, Chamberlain J, Robinson M, Moss S, Amar S, Balfour T et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. The Lancet. 1996;348(9040):1472-1477.
10. Kronborg O, Fenger C, Olsen J, Jørgensen O, Søndergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. The Lancet. 1996;348(9040):1467-1471.
11. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. Cochrane Database of Systematic Reviews. 2007;.

12. Steele R, McClements P, Libby G, Black R, Morton C, Birrell J et al. Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer. 2020.

13. Bowel screening - NICE CKS [Internet]. Cks.nice.org.uk. 2019 [cited 6 January 2020]. Available from: https://cks.nice.org.uk/bowel-screening

14. Bowel cancer statistics [Internet]. Cancer Research UK. 2019 [cited 6 January 2020]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer

15. Logan R, Patnick J, Nickerson C, Coleman L, Rutter M, von Wagner C. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. Gut. 2011;61(10):1439-1446.

16. Cancer | Scottish Bowel Screening Programme | Health Topics | ISD Scotland [Internet]. Isdscotland.org. 2019 [cited 6 January 2020]. Available from: https://www.isdscotland.org/Health-Topics/Cancer/Bowel-Screening/

17. Digby J, McDonald P, Strachan J, Libby G, Steele R, Fraser C. Use of a faecal immunochemical test narrows current gaps in uptake for sex, age and deprivation in a bowel cancer screening programme. Journal of Medical Screening. 2013;20(2):80-85.

18. Quyn A, Fraser C, Stanners G, Carey F, Carden C, Shaukat A et al. Uptake trends in the Scottish Bowel Screening Programme and the influences of age, sex, and deprivation. Journal of Medical Screening. 2017;25(1):24-31.

19. Steele R, McDonald P, Digby J, Brownlee L, Strachan J, Libby G et al. Clinical outcomes using a faecal immunochemical test for haemoglobin as a first-line test in a national programme constrained by colonoscopy capacity. United European Gastroenterology Journal. 2013;1(3):198-205.

20. Scottish Bowel Screening Statistics [Internet]. Isdscotland.org. 2019 [cited 6 January 2020]. Available from: https://www.isdscotland.org/Health-Topics/Cancer/Publications/2019-08-06/2019-08-06-Bowel-Screening-Publication-Report.pdf

21. Lansdorp-Vogelaar I, Knudsen A, Brenner H. Cost-effectiveness of Colorectal Cancer Screening. Epidemiologic Reviews. 2011;33(1):88-100.

22. Frazier A. Cost-effectiveness of Screening for Colorectal Cancer in the General Population. JAMA. 2000;284(15):1954.

23. Helm J, Russo M, Biddle A, Simpson K, Ransohoff D, Sandler R. Effectiveness and economic impact of screening for colorectal cancer by mass fecal occult blood testing. The American Journal of Gastroenterology. 2000;95(11):3250-3258.

24. Tappenden P, Chilcott J, Eggington S, Sakai H, Karnon J, Patnick J. Option appraisal of population-based colorectal cancer screening programmes in England. Gut. 2007;56(5):677-684.

25. Murphy J, Halloran S, Gray A. Cost-effectiveness of the faecal immunochemical test at a range of positivity thresholds compared with the guaiac faecal occult blood test in the NHS Bowel Cancer Screening Programme in England. BMJ Open. 2017;7(10):e017186.

26. Price J, Campbell C, Sells J, Weller D, Campbell H, Kenicer M et al. Impact of UK Colorectal Cancer Screening Pilot on hospital diagnostic services. Journal of Public Health. 2005;27(3):246-253.

27. KUMAR V. Robbins basic pathology. 9th ed. WB Saunders Elsevier; 2012.

28. Ponz de Leon M, Di Gregorio C. Pathology of colorectal cancer. Digestive and Liver Disease. 2001;33(4):372-388.

29. Muto T, Bussey H, Morson B. The evolution of cancer of the colon and rectum. Cancer. 1975;36(6):2251-2270.

30. Clark J, Collan Y, Eide T, Estève J, Ewen S, Gibbs N et al. Prevalence of polyps in an autopsy series from areas with varying incidence of large-bowel cancer. International Journal of Cancer. 1985;36(2):179-186.

31. SHINYA H, WOLFF W. Morphology, Anatomic Distribution and Cancer Potential of Colonic Polyps. Annals of Surgery. 1979;190(6):679-683.

32. Eide T. Remnants of adenomas in colorectal carcinomas. Cancer. 1983;51(10):1866-1872.

33. Pan J, Xin L, Ma Y, Hu L, Li Z. Colonoscopy Reduces Colorectal Cancer Incidence and Mortality in Patients With Non-Malignant Findings: A Meta-Analysis. American Journal of Gastroenterology. 2016;111(3):355-365.

34. Zauber A, Winawer S, O'Brien M, Lansdorp-Vogelaar I, van Ballegooijen M, Hankey B et al. Colonoscopic Polypectomy and Long-Term Prevention of Colorectal-Cancer Deaths. New England Journal of Medicine. 2012;366(8):687-696.

35. Winawer S, Zauber A, O'Brien M, Ho M, Gottlieb L, Sternberg S et al. Randomized Comparison of Surveillance Intervals after Colonoscopic Removal of Newly Diagnosed Adenomatous Polyps. New England Journal of Medicine. 1993;328(13):901-906.

36. Eide T. Risk of colorectal cancer in adenoma-bearing individuals within a defined population. International Journal of Cancer. 1986;38(2):173-176.

37. Cairns S, Scholefield J, Steele R, Dunlop M, Thomas H, Evans G et al. Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002). Gut. 2010;59(5):666-689.

38. Atkin W, Morson B, Cuzick J. Long-Term Risk of Colorectal Cancer after Excision of Rectosigmoid Adenomas. New England Journal of Medicine. 1992;326(10):658-662.

39. Farrar W, Sawhney M, Nelson D, Lederle F, Bond J. Colorectal Cancers Found After a Complete Colonoscopy. Clinical Gastroenterology and Hepatology. 2006;4(10):1259-1264.

40. Strum W. Colorectal Adenomas. New England Journal of Medicine. 2016;374(11):1065-1075.

41. Fearon E, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-767.

42. Deqin M, Chen Z, Nero C, Patel K, Daoud E, Cheng H et al. Somatic Deletions of the PolyA Tract in the 3′ Untranslated Region of Epidermal Growth Factor Receptor Are Common in Microsatellite Instability–High Endometrial and Colorectal Carcinomas. Archives of Pathology & Laboratory Medicine. 2012;136(5):510-516.

43. Tariq K, Tariq K, Ghias K, Ghias K. Colorectal cancer carcinogenesis: a review of mechanisms. Cancer Biology & Medicine. 2016;13(1):120-35.

44. Simons C, Hughes L, Smits K, Khalid-de Bakker C, de Bruïne A, Carvalho B et al. A novel classification of colorectal tumors based on microsatellite instability, the CpG island methylator phenotype and chromosomal instability: implications for prognosis. Annals of Oncology. 2013;24(8):2048-2056.

45. Colorectal Cancer Stages [Internet]. Cancer.org. 2018 [cited 3 March 2020]. Available from: https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/staged.html

46. Survival | Bowel cancer | Cancer Research UK [Internet]. Cancerresearchuk.org. 2018 [cited 3 March 2020]. Available from: https://www.cancerresearchuk.org/about-cancer/bowel-cancer/survival

47. Dukes' staging system | Bowel cancer | Cancer Research UK [Internet]. About-cancer.cancerresearchuk.org. 2018 [cited 3 March 2020]. Available from: https://about-cancer.cancerresearchuk.org/about-cancer/bowel-cancer/stages-types-and-grades/dukes-stagingfthe

48. Bowel cancer risk [Internet]. Cancer Research UK. 2019 [cited 6 January 2020]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/risk-factors#ref-0

49. Brown K, Rumgay H, Dunlop C, Ryan M, Quartly F, Cox A et al. The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015. British Journal of Cancer. 2018;118(8):1130-1141.

50. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective.

Continuous Update Project Expert Report 2018. Available at dietandcancerreport.org
51. Cross A, Sinha R. Meat-related mutagens/carcinogens in the etiology of colorectal cancer. Environmental and Molecular Mutagenesis. 2004;44(1):44-55.

52. Padmanabhan H, Brookes M, Iqbal T. Iron and colorectal cancer: evidence from in vitro and animal studies. Nutrition Reviews. 2015;73(5):308-317.

53. Santarelli R, Pierre F, Corpet D. Processed Meat and Colorectal Cancer: A Review of Epidemiologic and Experimental Evidence. Nutrition and Cancer. 2008;60(2):131-144.

54. Slavin J. Mechanisms for the Impact of Whole Grain Foods on Cancer Risk. Journal of the American College of Nutrition. 2000;19(sup3):300S-307S.

55. KIM K, TSAO R, YANG R, CUI S. Phenolic acid profiles and antioxidant activities of wheat bran extracts and the effect of hydrolysis conditions. Food Chemistry. 2006;95(3):466-473.

56. Slavin J. Why whole grains are protective: biological mechanisms. Proceedings of the Nutrition Society. 2003;62(1):129-134.

57. Bingham S. Mechanisms and experimental and epidemiological evidence relating dietary fibre (non-starch polysaccharides) and starch to protection against large bowel cancer. Proceedings of the Nutrition Society. 1990;49(2):153-171.

58. Seitz H, Stickel F. Molecular mechanisms of alcohol-mediated carcinogenesis. Nature Reviews Cancer. 2007;7(8):599-612.

59. Murphy N, Cross A, Abubakar M, Jenab M, Aleksandrova K, Boutron-Ruault M et al. A Nested Case–Control Study of Metabolically Defined Body Size Phenotypes and Risk of Colorectal Cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC). PLOS Medicine. 2016;13(4):e1001988.

60. Aleksandrova K, Pischon T, Buijsse B, May A, Peeters P, Bueno-de-Mesquita H et al. Adult weight change and risk of colorectal cancer in the European Prospective Investigation into Cancer and Nutrition. European Journal of Cancer. 2013;49(16):3526-3536.

61. Karahalios A, English D, Simpson J. Weight Change and Risk of Colorectal Cancer: A Systematic Review and Meta-Analysis. American Journal of Epidemiology. 2015;181(11):832-845.

62. Rapp K, Klenk J, Ulmer H, Concin H, Diem G, Oberaigner W et al. Weight change and cancer risk in a cohort of more than 65 000 adults in Austria. Annals of Oncology. 2008;19(4):641-648.

63. Song B, Cho K, Jo Y, Oh J, Kim Y. Colon Transit Time According to Physical Activity Level in Adults. Journal of Neurogastroenterology and Motility. 2012;18(1):64-69.

64. Liang P, Chen T, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: Systematic review and meta-analysis. International Journal of Cancer. 2009;124(10):2406-2415.

65. Kawachi I. A glossary for health inequalities. 2002.

66. Long Term Monitoring of Health Inequalities: Headline Indicators – October 2015 - gov.scot [Internet]. Gov.scot. 2015 [cited 3 March 2020]. Available from: https://www.gov.scot/publications/long-term-monitoring-health-inequalities-headline-indicators-october-2015/

67. Cancer by Deprivation in England 1996-2011 [Internet]. Ncin.org.uk. [cited 3 March 2020]. Available from: http://www.ncin.org.uk/about_ncin/cancer_by_deprivation_in_england

68. Cancer Awareness Measure (CAM) Key Findings Report; 2014 & Trends Analysis (2008-2014) [Internet]. Cancerresearchuk.org. 2016 [cited 3 March 2020]. Available from: https://www.cancerresearchuk.org/sites/default/files/cam_key_findings_report_-_2014_trends_analysis_v5.pdf

69. Doubeni C, Major J, Laiyemo A, Schootman M, Zauber A, Hollenbeck A et al. Contribution of Behavioral Risk Factors and Obesity to Socioeconomic Differences in Colorectal Cancer Incidence. JNCI Journal of the National Cancer Institute. 2012;104(18):1353-1362.

70. De Bruijn K, Arends L, Hansen B, Leeflang S, Ruiter R, van Eijck C. Systematic review and meta-analysis of the association between diabetes mellitus and incidence and mortality in breast and colorectal cancer. British Journal of Surgery. 2013;100(11):1421-1429.

71. Krämer H, Schöttker B, Raum E, Brenner H. Type 2 diabetes mellitus and colorectal cancer: Meta-analysis on sex-specific differences. European Journal of Cancer. 2012;48(9):1269-1282.

72. Jiang Y, Ben Q, Shen H, Lu W, Zhang Y, Zhu J. Diabetes mellitus and incidence and mortality of colorectal cancer: a systematic review and meta-analysis of cohort studies. European Journal of Epidemiology. 2011;26(11):863-876.

73. Larsson S, Orsini N, Wolk A. Diabetes Mellitus and Risk of Colorectal Cancer: A Meta-Analysis. JNCI: Journal of the National Cancer Institute. 2005;97(22):1679-1687.

74. Wu L, Yu C, Jiang H, Tang J, Huang H, Gao J et al. Diabetes Mellitus and the Occurrence of Colorectal Cancer: An Updated Meta-Analysis of Cohort Studies. Diabetes Technology & Therapeutics. 2013;15(5):419-427.

75. Singh S, Singh H, Singh P, Murad M, Limburg P. Antidiabetic Medications and the Risk of Colorectal Cancer in Patients with Diabetes Mellitus: A Systematic Review and Meta-analysis. Cancer Epidemiology Biomarkers & Prevention. 2013;22(12):2258-2268.

76. Lutgens M, van Oijen M, van der Heijden G, Vleggaar F, Siersema P, Oldenburg B. Declining Risk of Colorectal Cancer in Inflammatory Bowel Disease. Inflammatory Bowel Diseases. 2013;19(4):789-799.

77. Castaño-Milla C, Chaparro M, Gisbert J. Systematic review with meta-analysis: the declining risk of colorectal cancer in ulcerative colitis. Alimentary Pharmacology & Therapeutics. 2014;39(7):645-659.

78. Algra A, Rothwell P. Effects of regular aspirin on long-term cancer incidence and metastasis: a systematic comparison of evidence from observational studies versus randomised trials. The Lancet Oncology. 2012;13(5):518-527.

79. Cole B, Logan R, Halabi S, Benamouzig R, Sandler R, Grainge M et al. Aspirin for the Chemoprevention of Colorectal Adenomas: Meta-analysis of the Randomized Trials. JNCI: Journal of the National Cancer Institute. 2009;101(4):256-266.

80. Kim S, Baron J, Mott L, Burke C, Church T, McKeown-Eyssen G et al. Aspirin may be more effective in preventing colorectal adenomas in patients with higher BMI (United States). Cancer Causes & Control. 2006;17(10):1299-1304.

81. Lichtenstein P, Holm N, Verkasalo P, Iliadou A, Kaprio J, Koskenvuo M et al. Environmental and Heritable Factors in the Causation of Cancer — Analyses of Cohorts of Twins from Sweden, Denmark, and Finland. New England Journal of Medicine. 2000;343(2):78-85.

82. Grady W. Genetic testing for high-risk colon cancer patients1 1Abbreviations used in this paper: FAP, familial adenomatous polyposis; HMPS, hereditary mixed polyposis syndrome; HNPCC, hereditary nonpolyposis colon cancer; JPS, juvenile polyposis; MMR, mutation mismatch repair; MSI, microsatellite instability; PJS, Peutz-Jeghers syndrome; TGF, transforming growth factor. Gastroenterology. 2003;124(6):1574-1594.

83. Jasperson K, Tuohy T, Neklason D, Burt R. Hereditary and Familial Colon Cancer. Gastroenterology. 2010;138(6):2044-2058.

84. Armelao F, de Pretis G. Familial colorectal cancer: A review. World J Gastroenteroly [Internet]. 2014 [cited 3 March 2020];20 (28)(9292–9298). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4110560/

85. Jenkins M, Makalic E, Dowty J, Schmidt D, Dite G, MacInnis R et al. Quantifying the utility of single nucleotide polymorphisms to guide colorectal cancer screening. Future Oncology. 2016;12(4):503-513.

86. Haiman C, Le Marchand L, Yamamato J, Stram D, Sheng X, Kolonel L et al. A common genetic risk factor for colorectal and prostate cancer. Nature Genetics. 2007;39(8):954-956.

87. Tenesa A, Farrington S, Prendergast J, Porteous M, Walker M, Haq N et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nature Genetics. 2008;40(5):631-637.

88. Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, Spain S et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nature Genetics. 2007;39(8):984-988.

89. Theodoratou E, Montazeri Z, Hawken S, Allum G, Gong J, Tait V et al. Systematic Meta-Analyses and Field Synopsis of Genetic Association Studies in Colorectal Cancer. JNCI Journal of the National Cancer Institute. 2012;104(19):1433-1457.

90. Dove–Edwin I, de Jong A, Adams J, Mesher D, Lipton L, Sasieni P et al. Prospective Results of Surveillance Colonoscopy in Dominant Familial Colorectal Cancer With and Without Lynch Syndrome. Gastroenterology. 2006;130(7):1995-2000.

91. St. John D. Cancer Risk in Relatives of Patients with Common Colorectal Cancer. Annals of Internal Medicine. 1993;118(10):785.

92. Fuchs C, Giovannucci E, Colditz G, Hunter D, Speizer F, Willett W. A Prospective Study of Family History and the Risk of Colorectal Cancer. New England Journal of Medicine. 1994;331(25):1669-1674.

93. Roos V, Mangas-Sanjuan C, Rodriguez-Girondo M, Medina-Prado L, Steyerberg E, Bossuyt P et al. Effects of Family History on Relative and Absolute Risks for Colorectal Cancer: A Systematic Review and Meta-Analysis. Clinical Gastroenterology and Hepatology. 2019;17(13):2657-2667.e9.

94. Houlston R, Murday V, Harocopos C, Williams C, Slack J. Screening and genetic counselling for relatives of patients with colorectal cancer in a family cancer clinic. BMJ. 1990;301(6748):366-368.

95. Genetics of Colorectal Cancer (PDQ®)–Health Professional Version [Internet]. National Cancer Institute. 2020 [cited 4 March 2020]. Available from: https://www.cancer.gov/types/colorectal/hp/colorectal-genetics-pdq#link/_2986_toc

96. Lynch syndrome [Internet]. bowelcancerUK. [cited 4 March 2020]. Available from: https://www.bowelcanceruk.org.uk/campaigning/lynch-syndrome/

97. Lynch H, de la Chapelle A. Hereditary Colorectal Cancer. New England Journal of Medicine. 2003;348(10):919-932.

98. Rustgi A. The genetics of hereditary colon cancer. Genes & Development. 2007;21(20):2525-2538.

99. Barrow E, Robinson L, Alduaij W, Shenton A, Clancy T, Lalloo F et al. Cumulative lifetime incidence of extracolonic cancers in Lynch syndrome: a report of 121 families with proven mutations. Clinical Genetics. 2009;75(2):141-149.

100. Meyer L, Broaddus R, Lu K. Endometrial Cancer and Lynch Syndrome: Clinical and Pathologic Considerations. Cancer Control. 2009;16(1):14-22.

101. Sehgal R, Sheahan K, O'Connell P, Hanly A, Martin S, Winter D. Lynch Syndrome: An Updated Review. Genes. 2014;5(3):497-507.

102. Koornstra J, Mourits M, Sijmons R, Leliveld A, Hollema H, Kleibeuker J. Management of extracolonic tumours in patients with Lynch syndrome. The Lancet Oncology. 2009;10(4):400-408.

103. Kovacs M, Papp J, Szentirmay Z, Otto S, Olah E. Deletions removing the last exon of TACSTD1 constitute a distinct class of mutations predisposing to Lynch syndrome. Human Mutation. 2009;30(2):197-203.

104. Lynch syndrome [Internet]. Genetics Home Reference. 2020 [cited 4 March 2020]. Available from: https://ghr.nlm.nih.gov/condition/lynch-syndrome

105. Nojadeh J, Sharif S, Sakhinia E. Microsatellite instability in colorectal cancer. EXCLI [Internet]. 2018 [cited 4 March 2020];17. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938532/

106. Umar A, Boland C, Terdiman J, Syngal S, Chapelle A, Ruschoff J et al. Revised Bethesda Guidelines for Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndrome) and Microsatellite Instability. JNCI Journal of the National Cancer Institute. 2004;96(4):261-268.

107. EPCAM gene [Internet]. Genetics Home Reference. 2020 [cited 4 March 2020]. Available from: https://ghr.nlm.nih.gov/gene/EPCAM#conditions

108. Schneider R, Schneider C, Jakobeit C, Fürst A, Möslein G. Gender-Specific Aspects of Lynch Syndrome and Familial Adenomatous Polyposis. Viszeralmedizin. 2014;30(2):5-5.

109. Kohlmann W, Gruber S. Lynch Syndrome [Internet]. Ncbi.nlm.nih.gov. 2018 [cited 7 March 2020]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1211/

110. Zhang X, Li J. Era of universal testing of microsatellite instability in colorectal cancer. World Journal of Gastrointestinal Oncology. 2013;5(2):12.

111. Familial adenomatous polyposis [Internet]. Genetics Home Reference. 2020 [cited 4 March 2020]. Available from:

https://ghr.nlm.nih.gov/condition/familial-adenomatous-polyposis#statistics

112. Campbell W, Spence R, Parks T. Familial adenomatous polyposis. BJS. 1994;81(12):1722-1733.

113. Herrera L, O'Riordan B, Cambronero E. Familial Adenomatous Polyposis. Surgical Oncology Clinics of North America. 1994;3(3):419-434.

114. Petersen G, Slack J, Nakamura Y. Screening guidelines and premorbid diagnosis of familial adenomatous polyposis using linkage. Gastroenterology. 1991;100(6):1658-1664.

115. Vasen H, Moslein G, Alonso A, Aretz S, Bernstein I, Bertario L et al. Guidelines for the clinical management of familial adenomatous polyposis (FAP) [Internet]. BMJ. 2008 [cited 4 March 2020]. Available from: https://gut.bmj.com/content/57/5/704

116. APC gene [Internet]. Genetics Home Reference. 2020 [cited 7 March 2020]. Available from: https://ghr.nlm.nih.gov/gene/APC

117. Laurent-Puig P. APC gene: database of germline and somatic mutations in human tumors and cell lines. Nucleic Acids Research. 1998;26(1):269-270.

118. Spirio L, Olschwang S, Groden J, Robertson M, Samowitz W, Joslyn G et al. Alleles of the APC gene: An attenuated form of familial polyposis. Cell. 1993;75(5):951-957.

119. Yan H, Dobbie Z, Gruber S, Markowitz S, Romans K, Giardiello F et al. Small changes in expression affect predisposition to tumorigenesis. Nature Genetics. 2001;30(1):25-26.

120. Leppert M, Burt R, Hughes J, Samowitz W, Nakamura Y, Woodward S et al. Genetic Analysis of an Inherited Predisposition to Colon Cancer in a Family with a Variable Number of Adenomatous Polyps. New England Journal of Medicine. 1990;322(13):904-908.

121. Brensinger J, Laken S, Luce M, Powell S, Vance G, Ahnen D et al. Variable phenotype of familial adenomatous polyposis in pedigrees with 3′ mutation in the APC gene. Gut. 1998;43(4):548-552.

122. Giardiello F. Phenotypic Expression of Disease in Families That Have Mutations in the 5′ Region of the Adenomatous Polyposis Coli Gene. Annals of Internal Medicine. 1997;126(7):514.

123. Heppner Goss K, Trzepacz C, Tuohy T, Groden J. Attenuated APC alleles produce functional protein from internal translation initiation. Proceedings of the National Academy of Sciences. 2002;99(12):8161-8166.

124. Nieuwenhuis M, Vogt S, Jones N, Nielsen M, Hes F, Sampson J et al. Evidence for accelerated colorectal adenoma–carcinoma progression inMUTYH-associated polyposis?. Gut. 2011;61(5):734-738.

125. Hampel H. Genetic Testing for Hereditary Colorectal Cancer. Surgical Oncology Clinics of North America. 2009;18(4):687-703.

126. Sieber O, Lipton L, Crabtree M, Heinimann K, Fidalgo P, Phillips R et al. Multiple Colorectal Adenomas, Classic Adenomatous Polyposis, and Germ-Line Mutations inMYH. New England Journal of Medicine. 2003;348(9):791-799.

127. Jones N, Vogt S, Nielsen M, Christian D, Wark P, Eccles D et al. Increased Colorectal Cancer Incidence in Obligate Carriers of Heterozygous Mutations in MUTYH. Gastroenterology. 2009;137(2):489-494.e1.

128. Cleary S, Cotterchio M, Jenkins M, Kim H, Bristow R, Green R et al. Germline MutY Human Homologue Mutations and Colorectal Cancer: A Multisite Case-Control Study. Gastroenterology. 2009;136(4):1251-1260.

129. Nielsen M, Morreau H, Vasen H, Hes F. MUTYH-associated polyposis (MAP). Critical Reviews in Oncology/Hematology. 2011;79(1):1-16.

130. Boparai K, Dekker E, van Eeden S, Polak M, Bartelsman J, Mathus–Vliegen E et al. Hyperplastic Polyps and Sessile Serrated Adenomas as a Phenotypic Expression of MYH-Associated Polyposis. Gastroenterology. 2008;135(6):2014-2018.

131. Vogt S, Jones N, Christian D, Engel C, Nielsen M, Kaufmann A et al. Expanded Extracolonic Tumor Spectrum in MUTYH-Associated Polyposis. Gastroenterology. 2009;137(6):1976-1985.e10.

132. Lefevre J, Rodrigue C, Mourra N, Bennis M, Flejou J, Parc R et al. Implication of MYH in Colorectal Polyposis. Annals of Surgery. 2006;244(6):874-880.

133. MUTYH gene [Internet]. Genetics Home Reference. 2020 [cited 7 March 2020]. Available from: https://ghr.nlm.nih.gov/gene/MUTYH#location

134. Nielsen M, Joerink - van de Beld M, Jones N, Vogt S, Tops C, Vasen H et al. Analysis of MUTYH Genotypes and Colorectal Phenotypes in Patients With MUTYH-Associated Polyposis. Gastroenterology. 2009;136(2):471-476.

135. Knopperts A, Nielsen M, Niessen R, Tops C, Jorritsma B, Varkevisser J et al. Contribution of bi-allelic germline MUTYH mutations to early-onset and familial colorectal cancer and to low number of adenomatous polyps: case-series and literature review. Familial Cancer. 2012;12(1):43-50.

136. Sampson J, Dolwani S, Jones S, Eccles D, Ellis A, Evans D et al. Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. The Lancet. 2003;362(9377):39-41.

137. Latchford A, Neale K, Phillips R, Clark S. Juvenile Polyposis Syndrome. Diseases of the Colon & Rectum. 2012;55(10):1038-1043.

138. JASS J, WILLIAMS C, BUSSEY H, MORSON B. Juvenile polyposis-a precancerous condition. Histopathology. 2007;13(6):619-630.

139. Juvenile polyposis syndrome [Internet]. Genetics Home Reference. 2020 [cited 7 March 2020]. Available from: https://ghr.nlm.nih.gov/condition/juvenile-polyposis-syndrome

140. Gallione C, Repetto G, Legius E, Rustgi A, Schelley S, Tejpar S et al. A combined syndrome of juvenile polyposis and hereditary haemorrhagic telangiectasia associated with mutations in MADH4 (SMAD4). The Lancet. 2004;363(9412):852-859.

141. Dahdaleh F, Carr J, Calva D, Howe J. Juvenile polyposis and other intestinal polyposis syndromes with microdeletions of chromosome 10q22-23. Clinical Genetics. 2011;81(2):110-116.

142. SMAD4 gene [Internet]. Genetics Home Reference. 2020 [cited 7 March 2020]. Available from: https://ghr.nlm.nih.gov/gene/SMAD4

143. BMPR1A gene [Internet]. Genetics Home Reference. 2020 [cited 7 March 2020]. Available from: https://ghr.nlm.nih.gov/gene/BMPR1A#location

144. Zhou X, Woodford-Richens K, Lehtonen R, Kurose K, Aldred M, Hampel H et al. Germline Mutations in BMPR1A/ALK3 Cause a Subset of Cases of Juvenile Polyposis Syndrome and of Cowden and Bannayan-Riley-Ruvalcaba Syndromes*. The American Journal of Human Genetics. 2001;69(4):704-711.

145. Hereditary hemorrhagic telangiectasia [Internet]. Genetics Home Reference. 2020 [cited 7 March 2020]. Available from: https://ghr.nlm.nih.gov/condition/hereditary-hemorrhagic-telangiectasia#genes

146. Peutz-Jeghers syndrome [Internet]. Genetics Home Reference. 2020 [cited 7 March 2020]. Available from: https://ghr.nlm.nih.gov/condition/peutz-jeghers-syndrome#genes

147. Giardiello F, Brensinger J, Tersmette A, Goodman S, Petersen G, Booker S et al. Very high risk of cancer in familial Peutz–Jeghers syndrome. Gastroenterology. 2000;119(6):1447-1453.

148. STK11 gene [Internet]. Genetics Home Reference. 2020 [cited 7 March 2020]. Available from: https://ghr.nlm.nih.gov/gene/STK11#conditions

149. Nakau M, Miyoshi H, Seldin M, Imamura M, Oshima M, Taketo M. Hepatocellular Carcinoma Caused by Loss of Heterozygosity in Lkb1 Gene Knockout Mice. American Association for Cancer Research

[Internet]. 2005 [cited 7 March 2020];62(16). Available from: https://cancerres.aacrjournals.org/content/62/16/4549.article-info

150. Hearle N. Frequency and Spectrum of Cancers in the Peutz-Jeghers Syndrome. Clinical Cancer Research. 2006;12(10):3209-3215.

151. Vasen H, Mecklin J, Meera Khan P, Lynch H. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). Diseases of the Colon & Rectum. 1991;34(5):424-425.

152. VASEN H, WATSON P, MECKLIN J, LYNCH H. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative Group on HNPCC☆. Gastroenterology. 1999;116(6):1453-1456.

153. Peltomaki P, Vasen H. Mutations predisposing to hereditary nonpolyposis colorectal cancer: Database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. Gastroenterology. 1997;113(4):1146-1158.

154. Barnetson R, Tenesa A, Farrington S, Nicholl I, Cetnarskyj R, Porteous M et al. Identification and Survival of Carriers of Mutations in DNA Mismatch-Repair Genes in Colon Cancer. New England Journal of Medicine. 2006;354(26):2751-2763.

155. Wijnen J, Vasen H, Khan P, Zwinderman A, van der Klift H, Mulder A et al. Clinical Findings with Implications for Genetic Testing in Families with Clustering of Colorectal Cancer. New England Journal of Medicine. 1998;339(8):511-518.

156. Sepulveda A, Aisner D. Molecular Basis of Diseases of the Gastrointestinal Tract. Essential Concepts in Molecular Pathology. 2010;:243-261.

157. AGA technical review on hereditary colorectal cancer and genetic testing. Gastroenterology. 2001;121(1):198-213.

158. Terdiman J, Gum J, Conrad P, Miller G, Weinberg V, Crawley S et al. Efficient detection of hereditary nonpolyposis colorectal cancer gene carriers by screening for tumor microsatellite instability before germline genetic testing. Gastroenterology. 2001;120(1):21-30.

159. Syngal S. Sensitivity and specificity of clinical criteria for hereditary non-polyposis colorectal cancer associated mutations in MSH2 and MLH1. Journal of Medical Genetics. 2000;37(9):641-645.

160. Kastrinos F, Steyerberg E, Mercado R, Balmaña J, Holter S, Gallinger S et al. The PREMM1,2,6 Model Predicts Risk of MLH1, MSH2, and MSH6 Germline Mutations Based on Cancer History. Gastroenterology. 2011;140(1):73-81.e5.

161. PREMM 5 Model [Internet]. Premm.dfci.harvard.edu. 2020 [cited 7 March 2020]. Available from: https://premm.dfci.harvard.edu/premm/model_development

162. Kastrinos F, Uno H, Ukaegbu C, Alvero C, McFarland A, Yurgelun M et al. Development and Validation of the PREMM5 Model for Comprehensive Risk Assessment of Lynch Syndrome. Journal of Clinical Oncology. 2017;35(19):2165-2172.

163. Khan O, Blanco A, Conrad P, Gulden C, Moss T, Olopade O et al. Performance of Lynch Syndrome Predictive Models in a Multi-Center US Referral Population. American Journal of Gastroenterology. 2011;106(10):1822-1827.

164. Kastrinos F, Ojha R, Leenen C, Alvero C, Mercado R, Balmaña J et al. Comparison of Prediction Models for Lynch Syndrome Among Individuals With Colorectal Cancer. JNCI: Journal of the National Cancer Institute. 2015;108(2).

165. MMRpro [Internet]. Projects.iq.harvard.edu. 2020 [cited 7 March 2020]. Available from: https://projects.iq.harvard.edu/bayesmendel/mmrpro

166. MRC HGU - Prediction of DNA mismatch repair gene mutation status in incident colorectal cancer cases. [Internet]. Hnpccpredict.hgu.mrc.ac.uk. 2020 [cited 7 March 2020]. Available from: http://hnpccpredict.hgu.mrc.ac.uk

167. Weigl K, Chang-Claude J, Knebel P, Hsu L, Hoffmeister M, Brenner H. Strongly enhanced colorectal cancer risk stratification by combining family history and genetic risk score. Clinical Epidemiology. 2018;Volume 10:143-152.

168. Hsu L, Jeon J, Brenner H, Gruber S, Schoen R, Berndt S et al. A Model to Determine Colorectal Cancer Risk Using Common Genetic Susceptibility Loci. Gastroenterology. 2015;148(7):1330-1339.e14.

169. Jo J, Nam C, Sull J, Yun J, Kim S, Lee S et al. Prediction of Colorectal Cancer Risk Using a Genetic Risk Score: The Korean Cancer Prevention Study-II (KCPS-II). Genomics & Informatics. 2012;10(3):175.

170. Frampton M, Law P, Litchfield K, Morris E, Kerr D, Turnbull C et al. Implications of polygenic risk for personalised colorectal cancer screening. Annals of Oncology. 2016;27(3):429-434.

171. Mesher D, Dove-Edwin I, Sasieni P, Vasen H, Bernstein I, Royer-Pokora B et al. A pooled analysis of the outcome of prospective colonoscopic surveillance for familial colorectal cancer. International Journal of Cancer. 2013;134(4):939-947.

172. Dove-Edwin I, Sasieni P, Adams J, Thomas H. Prevention of colorectal cancer by colonoscopic surveillance in individuals with a family history of colorectal cancer: 16 year, prospective, follow-up study. BMJ. 2005;331(7524):1047.

173. Dunlop M, Campbell H. Screening for people with a family history of colorectal cancer. BMJ. 1997;314(7097):1779-1779.

174. Vasen H, Abdirahman M, Brohet R, Langers A, Kleibeuker J, van Kouwen M et al. One to 2-Year Surveillance Intervals Reduce Risk of Colorectal Cancer in Families With Lynch Syndrome. Gastroenterology. 2010;138(7):2300-2306.

175. Rex D, Boland R, Dominitz J, Giardiello F, Johnson D, Kaltenbach T et al. Colorectal Cancer Screening: Recommendations for Physicians and Patients from the U.S. Multi-Society Task Force on Colorectal Cancer. American Journal of Gastroenterology. 2017;112(7):1016-1030.

176. Diagnosis and management of colorectal cancer [Internet]. Sign.ac.uk. 2016 [cited 7 March 2020]. Available from: https://www.sign.ac.uk/assets/sign126.pdf

177. Engel C, Rahner N, Schulmann K, Holinski–Feder E, Goecke T, Schackert H et al. Efficacy of Annual Colonoscopic Surveillance in Individuals With Hereditary Nonpolyposis Colorectal Cancer. Clinical Gastroenterology and Hepatology. 2010;8(2):174-182.

178. Heiskanen I, Järvinen H. Fate of the rectal stump after colectomy and ileorectal anastomosis for familial adenomatous polyposis. International Journal of Colorectal Disease. 1997;12(1):9-13.

179. Bertario L, Russo A, Radice P, Varesco L, Eboli M, Spinelli P et al. Genotype and Phenotype Factors as Determinants for Rectal Stump Cancer in Patients With Familial Adenomatous Polyposis. Annals of Surgery. 2000;231(4):538-543.

180. Stjepanovic N, Moreira L, Carneiro F, Balaguer F, Cervantes A, Balmaña J et al. Hereditary gastrointestinal cancers: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. Annals of Oncology. 2019;30(10):1558-1571.

181. NCCN Flash Update: NCCN Guidelines® and NCCN Compendium® for Melanoma & NCCN Guidelines for Genetic/Familial High-Risk Assessment: Colorectal [Internet]. Nccn.org. 2018 [cited 7 March 2020]. Available from: https://www.nccn.org/about/news/ebulletin/ebulletindetail.aspx?ebulletinid=1470

182. Jones S. An introduction to power and sample size estimation. Emergency Medicine Journal. 2003;20(5):453-458.

183. Cancer | Cancer Statistics | Colorectal Cancer | Health Topics | ISD Scotland [Internet]. Isdscotland.org. [cited 14 March 2020]. Available from: https://www.isdscotland.org/Health-Topics/Cancer/Cancer-Statistics/Colorectal/

184. Murff H, Spigel D, Syngal S. Does This Patient Have a Family History of Cancer?. JAMA. 2004;292(12):1480.

185. Ziogas A, Anton-Culver H. Validation of family history data in cancer family registries. American Journal of Preventive Medicine. 2003;24(2):190-198.

186. Lindgren G. Adenoma prevalence and cancer risk in familial non-polyposis colorectal cancer. Gut. 2002;50(2):228-234.

Appendices:

Appendix 1: Caldicott approval letter for study

Information Governance
Maryfield House South
Mains Loan
Dundee
DD4 7BT
Tel. 01382 740074
Ext. 70249
www.nhstayside.scot.nhs.uk

Kyriaki Christou
Medical Student
138B Nethergate
Dundee
DD1 4ED

Date 25 October 2019
Your Ref
Our Ref IGTCAL5092+
Enquiries to Mr J.Donnelly
Extension 70249
Direct Line N/A
Email joseph.donnelly@nhs.net

Dear Kyriaki

CALDICOTT APPROVAL – Does Risk Assessment for Familial Colorectal Cancer in Tayside Work?

Proposal Sponsor: Dr Jonathon Berg, Honorary Consultant in Clinical Genetics, NHS Tayside

Data User(s): Kyriaki Christou, Medical Student, University of Dundee

Original Approval 27/09/18
Caldicott approval is given for you to access relevant and proportionate personal data in the Cancer Audit, Clinical Portal, ICE, and Clinical Pathology databases, and patient casenotes, for all patients with a diagnosis of familial colorectal cancer in Tayside between 2000-2010, in order to determine how successful risk stratification is in current Scottish clinical practice, as described in your application and supporting information.

It is noted that all patient identifiers will be securely saved separately and removed from the data immediately after use.

Extended Approval 25/10/19
Approval amended to extend cohort of patients to end of 2018 and to allow the data user to access the IMVU Genetic Database and UNISOFT and EMBASE endoscopic databases.

Thank you for your co-operation in providing us with the information requested by us in this process.
Please contact me should any queries arise from the application of this approval.

Yours sincerely

Joseph Donnelly

Joseph Donnelly
Information Governance Officer (Data Protection)
NHS Tayside

Copy to: Dr Jonathon Berg, Honorary Consultant in Clinical Genetics, NHS Tayside