GENETIC CONTRIBUTION TO COLORECTAL CANCER IN NEWFOUNDLAND AND LABRADOR BASED ON FAMILY HISTORY AND TUMOUR MOLECULAR ANALYSES FOR AN INCIDENT COHORT

AMANDA CRANE

GENETIC CONTRIBUTION TO COLORECTAL CANCER IN NEWFOUNDLAND AND LABRADOR BASED ON FAMILY HISTORY AND TUMOUR MOLECULAR ANALYSES FOR AN INCIDENT COHORT

by

© Dr. Amanda Crane

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the

requirements for the degree of

Master of Science in Medicine

Clinical Epidemiology

Department of Medicine

Memorial University of Newfoundland

May 2012

St. John's Newfoundland

Abstract

Background and Purpose: Colorectal cancer (CRC) is an insidious cancer associated with significant morbidity and mortality. In Canada, it is the second leading cause of cancer death. Newfoundland and Labrador has the highest rate of CRC of all ten Canadian provinces and is ideal for examining founder effects and studying incidence of hereditary disease in what is essentially a closed community. Family history is the strongest risk factor for this disease. Inherited deleterious variants in the mismatch repair (MMR) genes play a critical role in the development of some cases, but families exist who have a family history consistent with autosomal dominant inheritance and no known predisposing genes (Familial colorectal cancer type X, FCCTX). The aim of this study is to determine the genetic basis of CRC in Newfoundland and Labrador, defined by family history, molecular pathology of the CRC and molecular genetics, in an incident cohort of CRC cases occurring in the population during one year.

Methods: A study population of 262 consecutively diagnosed CRC cases from the Newfoundland population were identified and following application of the recruitment protocol 148 participants, 144 families, completed the study. Eligible probands were contacted for consent and with this permission tumour blocks were obtained. Detailed family histories were obtained and risk classified. Tests of microsatellite stability and immunohistochemistry for mismatch repair proteins in tumour blocks were performed. Cancer phenotypes in family members of patients with CRC were compared in various high risk groups.

Results: 12.5% of families (n=18) were classified as high risk according to Amsterdam criteria (AC) and Age and Cancer modified Amsterdam criteria (ACMAC). An additional 33.3% fulfilled the revised Bethesda guidelines for intermediate risk classification and 53.5% were low risk. Fifteen (10.9%) families demonstrated microsatellite instability with thirteen (86.7%) of these having a corresponding deficiency in mismatch repair proteins. Eleven of the fifteen high risk families (61.1%) were microsatellite stable (MSS) and had no mismatch repair protein deficiency and family history was consistent with FCCTX or age modified FCCTX. FCCTX family members had a lower frequency of both CRC and Lynch syndrome associated cancers as well as a later age of cancer onset compared to possible Lynch syndrome.

Conclusions: The incidence of CRC with high or intermediate risk family history is high in Newfoundland and Labrador. Families with a history consistent with autosomal dominant disease occurred in 12.5%; possible Lynch syndrome was suspected in a minority (35%) of these families, and FCCTX/age modified FCCTX in the majority. Phenotype differences exist between FCCTX and suspected LS families in both age of onset and cancer frequency among relatives. Opportunity for novel gene discovery exists because the molecular genetic basis for most of those with familial CRC was not determined.

Notes to the Reader

1. I started this thesis in 2002 after working as a research assistant with the Clinical Epidemiology Department. After completing the course work required for the masters program, I was accepted into medical school. Data collection and analysis were completed during medical school but due to the time restraints of being a medical resident, as well as being away in Kingston, Ontario, I was not able to dedicate the time which was required to complete my thesis. Upon conclusion of my residency and having moved back home to start a career in family medicine, I decided the completion of my Masters of Science was very important to me both personally and professionally. With the support of my supervisory committee and despite working full time I completed my thesis as it had been originally designed. Unfortunately, however, the results of this study could no longer be released as a preliminary analysis of a portion of the larger collaborative project as had been intended. The five year study had been completed and the results published. The detailed mutational analysis published by Woods et al.(2010) was never intended to be incorporated into this study as it was to be an extension of and more inclusive than these results. It was completed by a separate team of researchers than those whom I received the molecular data results. All of my data had been compiled and analyzed prior to the release of this study but where possible I have made reference to corresponding results published by Woods et al. (2010). Significant advancements have been made in the research of Lynch syndrome since the collection of this data and I have attempted to present my data in light of these. While mutation analysis is the gold standard in diagnosis of LS,

it is also very expensive. With a more detailed analysis of family history risk classification along with the more economical molecular analyses, the clinical relevance of the information presented in this thesis should not be underestimated.

2. This project involves the incident CRC cases in the first year of data collected for a CIHR funded 5 year study. My two year involvement with this project initially began as a research assistant and then as a masters student. My main research role included helping with recruitment of participants, collection of the family history data and risk classification of families recruited for the entire project. For the population subset included in this thesis, I was totally responsible for the family history risk classification as well as the analysis of the molecular information in combination with family history data. Details of the work performed as well as identifying that contributed by other team members can be found throughout the methods section.

3. Since the start of my work, the terminology related to familial CRC has evolved. Most in the field now use the term **Lynch syndrome** (LS) as the preferred synonym over Hereditary Non-Polyposis Colorectal Cancer (HNPCC) and is used as such throughout the thesis. Specifically, LS is the genetic definition indicating a patient has an inherited pathogenic variant in an MMR gene while HNPCC is the clinical definition based on one's family history.

4. The molecular component of this research was completed and provided by Dr. Roger Green and the molecular lab teams. Without their work this thesis would not have been possible.

Acknowledgments

I would like to thank my supervisor Dr. Patrick Parfrey and supervisory committee for providing me the opportunity to complete this project as well as for their support despite the interruption of medical school and residency training; Dr. Roger Green and the molecular lab teams for providing me with the molecular results; Dr. Jane Green for her expertise in pedigree analysis and family history classification; Dr. Elizabeth Dicks and the recruitment team for their continued support and participant information; My family and friends for their encouragement over these last 9 years. Finally, I would like to thank my husband Stephen Crane who brought me back home and stood by me throughout this long process.

Table of Contents

List of Tables

List of Figures

Abbreviations

ACMAC	Age and Cancer Modified Amsterdam Criteria
AFAP	Attenuated Familial Adenomatous Polyposis
AC I	Amsterdam Criteria I
ACII	Amsterdam Criteria II
APC	Adenomatous Polyposis Coli gene
CRC	Colorectal Cancer
CRC-IHRT	Colorectal Cancer Interdisciplinary Health Research Team
FAP	Familial Adenomatous Polyposis
FCCTX	Familial Colorectal Cancer Type X
FHQ	Family History Questionnaire
HNPCC	Hereditary Nonpolyposis Colorectal Cancer
ICG-HNPCC	International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer
IHC	Immunohistochemistry
IR	Intermediate Risk
LS	Lynch Syndrome
MAP	MUTYH-Associated Polyposis
MMR	Mismatch Repair
MSI	Microsatellite Instability
MSS	Microsatellite Stable
NL	Newfoundland and Labrador
PMGP	Provincial Medical Genetics Program
RBG	Revised Bethesda Guidelines
ROI	Request of Information

Chapter 1: *Introduction*

1.1 *Colorectal Cancer*

1.1.1 *Definition*

Colorectal cancer (CRC) is an insidious cancer originating in the colon or rectum. In Canada, it is the second leading cancer causing death for men and women combined. An estimated 22,200 Canadians will be diagnosed with colorectal cancer in 2011 with 8,900 dying from their disease. When colorectal cancer is diagnosed in its early stages, it is curable. This emphasizes the importance and value in effective and appropriate screening for those at increased risk (Canadian Cancer Society, July 16, 2011).

Colorectal cancer can be broadly classified into two main categories: inherited and sporadic. Up to thirty percent of CRC has a hereditary component (Lichtenstein et al., 2000; Grady, 2003). Mendelian inherited conditions account for 5% of cases and have a recognized phenotype. The remaining cases with a family history of CRC are less well defined (Lynch & de la Chapelle, 2003).

Well defined inherited forms of colorectal cancer are based on clinical, pathological and genetic criteria. The most common hereditary forms include Lynch syndrome (LS), familial adenomatous polyposis (FAP), and MUTYH-Associated polyposis (MAP).

1.1.2 *HNPCC to Lynch Syndrome*

Throughout the decades many names have been used to classify this most common hereditary form of CRC. Initially, it was called family cancer syndrome which, in the 1980's, was changed to Lynch syndrome I and Lynch syndrome II, differentiating those that had primarily colorectal cancer from those who also had uterine and other extra-colonic cancers (Boland & Truncale, 1984). In order to distinguish this syndrome from FAP the term hereditary nonpolyposis hereditary colon cancer (HNPCC) was coined (Lynch et al., 1985). HNPCC has no pathognomonic feature and cancer is frequently found at extra-colonic sites. Multiple classification systems including Amsterdam criteria I (ACI), Amsterdam criteria II (ACII) and Bethesda guidelines have been developed and used in clinical practice to identify individuals at risk for HNPCC who require further evaluation. However, since the molecular elucidation of the cause of HNPCC and the role of mismatch repair (MMR) genes, it has been proposed that families having a known genetic variant in a MMR gene be referred to as LS while those with no molecular diagnosis yet fulfilling the ACI be referred to as Familial Colorectal Cancer-type X (FCCTX) (Lindor et al., 2005; Boland, 2005).

1.1.3 *FAP*

This syndrome is the second most common inherited form of CRC with a prevalence of 1 in 10,000. Its characteristic feature is the presence of hundreds to thousands of colonic polyps. This clinical phenotype allows for an easier diagnosis of this hereditary disorder. An attenuated form (AFAP) has also been described which has a

variable number of polyps and reduced penetrance. Both syndromes are attributed to deleterious germline variations of the adenomatous polyposis coli (*APC*) gene (Lynch & de la Chapelle, 2003).

1.1.4 *MAP*

MAP is characterized by polyposis of the colorectum as well as an increased risk of CRC. It is an autosomal recessive inherited syndrome which means siblings of an affected patient have a 25% chance of being affected and 75% chance of being a carrier of the associated variant. It is caused by biallelic mutations in *MUTYH* and genetic testing for MAP should be considered in those with greater than 10 colorectal adenomas but do not have an identifiable mutation in *APC* (Jasperson et al., 2010).

1.2 *Nomenclature of Hereditary Colon Cancer*

Hereditary colon cancer can be difficult to recognize and diagnose without established clinical criteria. In addition, as research continues, a framework of classification is imperative to appropriately compare results and make scientific advances. As already indicated, HNPCC was the initial nomenclature indicating high familial risk prior to molecular and genetic advancements leading to the classifications of LS and FCCTX.

1.2.1 *Amsterdam Criteria I*

In 1990, to help provide uniformity for collaborative studies, the International

Collaborative Group on HNPCC (ICG-HNPCC) established a set of criteria to define LS families:

> There should be at least 3 relatives with CRC and all the following criteria should be present:
>
> - One should be a first degree relative of the other 2
> - At least 2 successive generations should be affected
> - At least 1 CRC should be diagnosed before age 50
> - Familial adenomatous polyposis should be excluded
> - Tumours should be verified by pathological examination
>
> (Vasen, 1991)

Although these criteria provided uniformity for collaborative studies they were very restrictive. Researchers believed many families would be excluded as extra-colonic cancers known to be part of the syndrome were not included. Consequently, families at risk would not receive appropriate hereditary designation and clinically would not avail of necessary screening and follow up (Vasen, 1999).

1.2.2 *Amsterdam Criteria II*

Established in 1997, this second set of criteria emerged aiming to provide simple, clinical criteria that had a high likelihood of identifying LS families. The goal was to enable a clinical diagnosis without molecular confirmation as not all families would have access to genetic analyses. The important feature, excluded in the original criteria, is the

inclusion of extra-colonic tumours. Studies have found cancer of the endometrium, stomach, ovaries, renal pelvis, brain, and hepatobiliary tract are the most highly associated with LS. Of these, cancer of the endometrium, ureter, renal pelvis, and small bowel are considered the most specific and therefore included in the Amsterdam II criteria (Watson & Lynch, 1993; Vasen et al., 1996; Sijmons, 1998; Aarnio et al., 1995):

> There should be at least 3 relatives with an LS-associated cancer (CRC, cancer of the endometrium, small bowel, ureter, or renal pelvis):
>
> - One should be a first degree relative of the other 2
> - At least 2 successive generations should be affected
> - At least 1 should be diagnosed before age 50
> - Familial adenomatous polyposis should be excluded
> - Tumours should be verified by pathological examination

1.2.3 *Molecular Characterization*

The molecular basis of LS has been established for nearly two decades. A deleterious variant in a DNA MMR gene (*MLH1*, *MSH2*, *MSH6*, or *PMS2*) is responsible for LS. The MMR system maintains genetic stability through correction of single base pair mismatches and insertion-deletion loops that form during DNA replication. Some variants in these genes result in an absence of mismatch repair and an accumulation of errors in stretches of sequences known as microsatellites. Microsatellites are short segments of DNA that have a repeated sequence and are particularly prone to insertion/deletion variants. In the absence of repair, the length of the microsatellite is

altered which is known as microsatellite instability (MSI). Colorectal cancers develop slowly through the acquisition of these genetic alterations that cannot be corrected by the repair mechanisms built into each cell. There are only a few common variants but each tumour has its own genetic alteration making the task of planning treatment very difficult (Boland & Goel, 2010). Early identification of a hereditary colorectal cancer syndrome allows appropriate surveillance, and early detection and treatment for family members leading to reduced morbidity and mortality in carriers (Jarvinen et al., 2000; Wagner et al., 2005). Through molecular advancements, these families can be more accurately identified through tumor and blood genetic testing.

1.2.4 *Bethesda Guidelines*

With a better understanding of the clinical and pathological manifestations of LS a new set of criteria were developed to identify tumours that should undergo molecular tumour analyses including microsatellite instability and/or immunohistochemical staining. If the results of these tests implicated LS, it was recommended to carry out germline testing of DNA MMR genes to confirm the diagnosis. In contrast to AC I and II, which focussed primarily on an extended family history of CRC and LS-associated tumours, the Bethesda guidelines incorporate young onset and/or multiple primary tumours in patients whose family history is either incomplete or the pedigree is small, as well as it includes histopathological presentations. The criteria initially proposed in 1996, and revised in 2002, have been demonstrated to be valuable in selecting families for DNA screening (Vasen et al., 2007; Julie et al., 2008).

Classical and revised Bethesda guidelines (RBG) include:

Classical Guidelines:

- Individuals with two LS-related cancers, including synchronous and metachronous colorectal cancers or associated extra-colonic cancers (endometrial, ovarian, gastric, hepatobiliary, small bowel cancer or transitional cell carcinoma of the renal pelvis or ureter).
- Individuals with colorectal cancer or endometrial cancer and a first degree relative with colorectal cancer and/or LS-related extra-colonic cancer and/or colorectal adenoma; one of the cancers diagnosed at age < 45 years, and the adenoma diagnosed at age < 40 years.
- Individuals with right-sided colorectal cancer or endometrial cancer diagnosed at age < 45 years.
- Individuals with right-sided colorectal cancer with an undifferentiated pattern on histopathology diagnosed at age < 45 years.
- Individuals with signet-ring-cell-type colorectal cancer diagnosed at age < 45 years.
- Individuals with adenomas diagnosed at age < 45 years.

Revised Bethesda Guidelines:

- Colorectal cancer diagnosed in a patient < 50 years old.
- Presence of synchronous, metachronous colorectal, or other LS-associated tumours (colorectal, endometrial, stomach, ovarian, pancreas, ureter, renal

pelvis, biliary tract, and brain tumours, sebaceous gland adenomas and keratoacanthomas in Muir Torre syndrome and carcinoma of the small bowel) regardless of age.

- Colorectal cancer with high microsatellite instability (MSI-H) histology* diagnosed in a patient < 60 years of age.
- Individual with colorectal cancer and one or more first-degree relatives with a LS-related tumor, with one of the cancers diagnosed under age of 50 years.
- Individual with colorectal cancer and two or more first-degree or second-degree relatives with LS-related tumours, regardless of age.

---

*Presence of tumour infiltrating lymphocytes, Crohn's-like lymphocytic reaction, mucinous/signet-ring differentiation, or medullary growth pattern

(Umar et al., 2004)

1.2.5 *Lynch Syndrome*

After multiple international conferences and molecular advancements in the study of HNPCC, a consensus to more specifically define the term LS was reached. Familial colon cancer clustering was to be split between those caused by variants in the *APC* gene causing FAP or those caused by failure of the DNA MMR system (Boland, 2005). Familial clusters whose tumours demonstrate failure of the MMR system would be referred to as possible LS while confirmation of diagnosis requires germline testing of DNA MMR genes (Vasen et al., 2007; Julie et al., 2008). Interestingly, many individuals with tumours demonstrating a mutated DNA MMR gene do not meet the Amsterdam criteria for family history. Additionally, only about 40% of those that meet the

Amsterdam criteria have evidence of a mutated DNA mismatch repair gene (Lindor et al., 2005).

1.2.6 *Familial Colorectal Cancer Type X*

The subset of families (60%) fulfilling ACI but not demonstrating MMR deficiency or an inherited MMR variant have been described as having FCCTX. A separate classification for these families was deemed necessary following research demonstrating important differences. Lindor et al. (2005) compared 90 ACI families with MMR deficiency with 71 ACI families without MMR defects. Proband tumour MSI was used to assign MMR status. They demonstrated MMR deficient families had a statistically significant increased risk of developing colorectal, endometrial, gastric, small intestine, and kidney cancer as expected for LS. The MMR proficient families, however, did not have an increased risk of cancer at any site. Additionally, the average age of diagnosis of CRC was later (age 61) in MMR proficient families compared to those with deficiency whose average age was 49 years old.

While the underlying mechanisms of carcinogenesis have not yet been elucidated, it is hypothesized that some familial aggregation may occur by chance, some by shared lifestyle factors, and some by a novel genetic mechanism(s) (Lindor et al., 2005; Kerber et al., 2005). Consequently, FCCTX, which is likely caused by multiple genetic insults, is recognized as a separate clinical category with a lower incidence of extra-colonic cancers and later age of onset of colorectal cancer (Mueller-Koch et al., 2005).

1.2.7 *Common Familial Risk Colon Cancer*

A simple family history of CRC has been associated with an increased familial risk. In fact, those who have a first degree relative with CRC diagnosed after the age of 50 have a two to three times increased risk for CRC. Population based studies have estimated 20% of all CRC occur in a higher risk setting defined by the occurrence of CRC before age 50, or a first degree relative pair with CRC (Kerber et al., 2005). The risk associated with family history varies greatly according to the age of onset of CRC in the family members, the number of affected relatives, the closeness of the genetic relationship, and whether cancers have occurred across generations (Fuchs et al., 1994; St.John et al., 1993). The presence of CRC in more than one family member may be caused by shared environmental risk factors or even chance but there is growing research implicating a hereditary component. Common familial CRC is believed to result from a number of lower penetrant susceptibility genes other than those associated with defined hereditary syndromes (Jasperson et al., 2010).

1.3 *Clinical Features*

LS is inherited in an autosomal dominant pattern. There is a fifty percent risk that each child of a carrier will inherit the variant and have a high risk of developing colorectal or other associated extra-colonic cancers. Polyps and colorectal cancers are predominately in the right side of the colon and there is a small and finite number, unlike in FAP which may have hundreds to thousands of colonic polyps. Cancer usually develops at a younger age than that seen in the general population and synchronous and

metachronous tumours are common (Hampel et al., 2005). Histologically, tumours are often poorly differentiated, demonstrate a medullary growth pattern, heavy infiltration of lymphocytes, or a mucinous or signet ring cell differentiation (Lindor et al., 2006).

1.4 *Colorectal Cancer and Family History*

As already identified, a key determinant in the diagnosis of a hereditary cancer syndrome is a detailed review of the family history. Valuable details include all cancer types and sites, age at onset, existence of multiple primaries, and any pathological findings (Lynch & de la Chapelle, 2005). A pedigree of family members can then be constructed and evaluated according to the established criteria defining the hereditary cancer syndromes. Studies evaluating the accuracy of reported clinical data for families meeting ACI or ACII have shown to be reliable. Love et al. (1985) demonstrated confirmation of 39 out of 42 reported colonic cancers among first degree relatives. Sijmons et al. (2000) reported 89% accuracy of reported cancers in 120 families and Kerber et al. (1998) confirmed 11 out 17 reported colonic cancers in first degree relatives. However, not all studies have such proven accuracy and therefore, as recommended by Katballe et al. (2001), verification of reported cancers and further exploration of family members should be standard procedure. A hereditary cancer syndrome should not be rejected simply because ACI or ACII are not fulfilled based on family history alone. Additionally, basing a diagnosis of hereditary cancer solely on the phenotypical analysis of the family pedigree can lead to both under- and overestimation of the frequency of the disease. Small pedigrees should be expanded to reliably exclude or include possible LS,

and pedigrees of adequate size are needed to eliminate chance aggregation (Percesepe et al., 1995; Katballe et al., 2001).

1.5 *Microsatellite Instability*

Microsatellites are sequences of repetitive DNA with repeating units of 1 to 7 base pairs. These regions are prone to replication errors, which are normally repaired by MMR enzymes. If one of the MMR enzymes is impaired, such as in LS, the microsatellites become unstable and will have varying repeat numbers at specific loci. MSI is defined as any change in length of a microsatellite due to insertion or deletion of repeating units.

Diagnostic criteria (RBG) including a reference panel of five markers were developed in 1998 to provide a uniform tool to identify and characterize microsatellite unstable tumours (Boland et al., 1998). Initially, MSI was classified as low or high. Tumours were MSI-high if two or more markers were unstable and MSI-low if only one marker was unstable. Tumours were microsatellite stable (MSS) in the absence of any unstable markers (Boland et al., 1998). Much debate has surrounded any clinical significance between MSI-low and MSS. A review of the literature conducted by de la Chapelle and Hampel (2010) concluded there was no convincing molecular or physiologic difference between MSI-low and MSS tumours. Consequently, they recommended MSI-high be referred to as simply MSI and MSI-low as MSS.

Microsatellite instability is generally a phenotypic marker of defective MMR involving proteins *MLH1*, *MSH2*, *MSH6*, and *PMS2*. While it identifies the presence of defective MMR it does not identify which protein/gene is affected. MSI has been

considered the hallmark of cancers found in LS but has also been observed in 15-20% of sporadic colon cancers (Boland, 2007; Poulogiannis et al., 2010). Consequently, most CRCs with MSI are not caused by LS. In sporadic colon cancers, this phenotype is almost entirely due to inactivation of the *MLH1* gene. This occurs due to hypermethylation of the promoter region resulting in silencing both copies of *MLH1* (Kane et al., 1997; Poulogiannis et al,. 2010). This acquired silencing is usually found in older patients but is occasionally also found in younger patients (Gryfe, 2006). About 75-80% of the CRCs which show MSI (about 10-12% of all CRCs) develop instability due to this acquired defect and the other 20-25% (3-4% of all CRCs) are due to LS (Boland et al., 2008). The characteristic features of sporadic CRC with MSI include the absence of significant familial clustering, methylation of the *MLH1* promoter, and absence of *MLH1* and *PSM2* proteins (Boland & Goel, 2010). In LS, there is a corresponding germline deleterious variant in a mismatch repair gene including *MLH1*, *MSH2*, *MSH6*, *PMS2*, and very rarely *PMS1*. Identification of MSI is useful to identify possible LS families, to study genetic mechanisms of colorectal carcinogenesis, and for grouping patients in treatment trials and prognostic marker studies (Esemuede et al., 2010). It is a valuable tool to identify LS when families do not meet the clinical criteria and would otherwise be missed.

Importantly, there is an absence of MSI in a small number of LS tumours. It is believed this is either due to a false negative secondary to an inadequate number of markers or tumour cells in the sample, or it is a phenocopy and hence a sporadic tumour in an individual with LS. Practically speaking, it was recommended that without a strong

family history, tumours that are MSS from the five-marker Bethesda panel need no further evaluation for possible LS or somatic inactivation of a MMR gene (de la Chapelle & Hampel, 2010). Despite this fact, 5-20% of LS tumours do not exhibit any MMR protein deficiency on IHC analysis even though they have lost MMR function as demonstrated by MSI when testing tumour DNA. While the cause of this discrepancy remains unknown it could be due to a yet to be identified MMR protein (Poulogiannis et al., 2010).

Three clinical uses have been outlined for the use of MSI status in CRC patients. Firstly, it is used to identify possible LS, secondly, survival is significantly better in young patients with MSI and thirdly, these tumours do not have the same response to chemotherapy as MSS tumours (Boland et al., 2008).

1.6 *DNA MMR System*

This system is composed of several proteins that work together to detect and repair errors during DNA replication. The two families of proteins involved are the MutS homologue (MSH), signaling the site of mispairing, and MutL homologue (MLH), needed to complete the repair process. Functioning as heterodimers, *MSH2* is an obligatory partner often pairing with *MSH6* or *MSH3*. In the MutL family, *MLH1* is the obligatory protein which may pair with *PMS2*, *PMS1* or *MSH3* (Boland et al., 2008). In summary, *MSH6* can only pair with *MSH2* and *PMS2* can only pair with *MLH1*. Therefore, when *MSH2* is inactivated, IHC staining will usually show absence of both *MSH2* and *MSH6* proteins, whereas, when *MSH6* is lost, *MSH2* staining remains positive.

Similarly, staining loss of *MLH1* and *PMS2* demonstrates inactivation of *MLH1* whereas only loss of *PSM2* equates to a mutated *PMS2* (see Table A below).

Table A: Loss of Immunohistochemistry Staining in Relation to Affected Mismatch Repair Gene

	Affected Gene			
Immunohistochemical Loss of Staining	*MSH2*	*MSH6*	*MLH1*	*PSM2*
MSH2 and *MSH6*	+	Occasional		
MSH6		+		
MLH1 and *PMS2*			+	Rare
PMS2				+

(de la Chapelle & Hampel, 2010)

Additionally, MSI activity is impacted by the particular protein absent. If *MSH2* is absent there is no MMR activity and typical MSI is demonstrated. However, if *MSH6* is the only protein absent, some DNA MMR is preserved and typical MSI may not be observed (Kolodner et al., 1999).

Demonstrating a deleterious germline variant is the gold standard for diagnosing LS, however, this is time consuming and expensive. Immunohistochemistry (IHC) analysis, on the other hand, with/without MSI is inexpensive and is also a predictor of LS (Trano et al., 2010).

1.6.1 *MLH1 Promoter Methylation*

Additional mechanisms for DNA MMR gene inactivation were sought following the discovery of MSI phenotype in 15% of colorectal cancers from patients lacking a

significant family history (Herman, 1998). Widespread changes in DNA methylation had long been observed in colon cancer leading to the discovery that MMR activity could be impaired not only by germline variants in MMR but also by somatic variants and epigenetic modifications such as DNA methylation which results in gene silencing (Herman et al., 1998; Veigl et al., 1998). In a study by Thibodeau at al. (1998) 40 of 42 (95%) sporadic MSI and MMR deficient tumours lacked expression of *MLH1* indicating *MLH1* had a principal role in the phenotype of sporadic MSI CRCs. Subsequent studies were able to demonstrate aberrant promoter *MLH1*methylation was the likely explanation for the loss of *MLH1* in these sporadic tumours, and approximately 80-90% of sporadic MSI CRCs exhibit defective MMR function due to *MLH1* promoter methylation (Herman et al., 1998; Veigl et al., 1998). *MLH1* methylation is functionally equivalent to an inactivating mutation and produces MSI in tumours not associated with LS. Consequently, while isolated MSI cannot confirm LS, methylation in combination with MSI could be useful for discriminating sporadic from LS-associated CRC (Hitchins et al., 2007).

Interestingly, *MLH1* silencing may no longer be considered a mechanism exclusive to sporadic tumours. A small proportion of LS tumours have also been found to demonstrate *MLH1* promoter hypermethylation. Germline methylation was first reported by Suter et al. (2004) in two individuals who met clinical criteria for LS but no MMR gene mutation could be found.

1.7 *Correlation between MSI and MMR*

Several studies have examined the concordance between MSI and MMR. In one study by Mueller et al. (2009) which analysed seventy one CRC cases suspected to be LS based on family history, twenty-eight were MSI with 96% of these linked to a defect in a known MMR gene. Overall, consensus in the literature indicates both MSI testing and the four antibody IHC staining (*MLH1*, *MSH2*, *MSH6*, and *PMS2*) are fairly equivalent in predicting germline variants, although the concordance is not perfect. The sensitivity of IHC predicting MSI has been reported to be about 92% as both tests may miss cases detectable by the other. For example, IHC can detect *MSH6* variant cases that may not show a high frequency MSI and/or MSI may detect cases that are not covered by the IHC panel (Shia et al., 2004). Additionally, 5-20% of Lynch related tumours do not exhibit any detectable MMR protein expression abnormality despite demonstrating MSI (Poulogiannis et al., 2010).

1.8 *Family History and Molecular Testing*

Amsterdam criteria are currently used to identify possible LS patients based on family history while Bethesda guidelines identify patients who should undergo tumour molecular analyses for diagnosis. Consequently, if either family history or tumour analyses are positive, germline testing of DNA MMR genes should be completed to confirm the diagnosis. (Vasen et al., 2007; Julie et al., 2008).

In a prospective study by Trano et al. (2010), the performance of clinical guidelines and molecular tumour tests (microsatellite instability, BRAF mutation, and methylation of

MMR genes) were assessed to identify possible LS. Three hundred thirty-six tumours from unselected, consecutively diagnosed CRC patients were analyzed by molecular tumour tests and the patients were classified according to RGB and ACII. A total of eighty-seven patients fulfilled the RBG for molecular tumour analyses and eight fulfilled the ACII. Molecular testing identified twelve tumours as probable LS. RBG had identified six of them and therefore likely LS, while five of the eight meeting ACII were not likely LS as they did not fulfill the molecular tumour criteria. They also found that half of the tumours that were highly suspicious of LS did not fulfill the RBG.

Another study completed by Syngal et al. (2000) classified seventy families by clinical criteria and completed mutation analysis of *MSH2* and *MLH1* by full gene sequencing. They found the sensitivity and specificity of the ACI were 61% and 67%. The sensitivity and specificity of the ACII were 72% and 78%, respectively. The Bethesda guidelines were the most sensitive with a sensitivity of 94% and specificity of 25%. These results considered only definitive pathogenic variants. In addition, investigators acknowledged that 40% of families fulfilling Amsterdam criteria did not have *MSH2* or *MLH1* deleterious variants and thus additional genes remain to be discovered. In a study by Valle et al. (2007), the clinical utility of MSI to accurately classify familial clusters of CRC was demonstrated. Only ACI families were identified from a Spanish national registry and tumours tested. Of those studied, 59.4% were MSI and classified as LS while the remaining 40.6% were MSS and classified as FCCTX. MSI families had earlier onset cancer, proximally located tumour, and a tumour spectrum consistent with LS, whereas most MSS families had only CRC. Additionally, in a study

by Perea et al. (2010), MSI was demonstrated in 31% of tumours from young patients. Clinicopathological features included earlier age of onset, signet-ring cell tumours, more proximally located tumours, and poorly differentiated tumours with higher mucin production suggesting LS.

Consequently, distinguishing LS from sporadic CRC requires consideration of a wide range of evidence: from the family history and examination of the tumours for MSI and abnormal MMR protein expression, to germ-line mutation analysis (Poulogiannis et al., 2010).

1.9 *Incidence*

Incidence of familial CRC has been estimated through the use of family history and/or molecular characterization. As indicated above, criteria, such as the Amsterdam criteria, have been developed to provide a scientific basis of comparison across studies. Additionally, for effective cancer prevention and research it is imperative to have an accurate knowledge of the cancer frequency.

Numerous studies have been completed estimating the incidence of familial CRC with the most accurate stemming from the population based studies. By eliminating selection bias and providing an unselected patient population these studies provide more reliable estimates of incidence. Using the Amsterdam criteria a variety of population based estimates ranging from 0.3-5.8% have been reported. In a large, multi-centre study in Finland by Mecklin et al. (1995), 0.7-2.4% of CRC was estimated to be familial. A smaller study, also by Mecklin (1987), which included only one Finnish province,

demonstrated a higher estimate of 3.5-5.5% of all CRC was familial. Similarly, Ponz de Leon (1993) noted a 3.4-4.5% incidence in Northern Italy. The lowest estimate was found by Aaltonen et al. (1994) who linked cancer registry with population registry data and produced an estimate of 0.5-0.9%.

Other attempts at estimation included extrapolation to the larger population from its incidence in younger patients. Kee and Collins (1991) estimated 1-2.6% would be familial from an incidence of 6% found in patients less than fifty five years of age. Finally, Westlake et al. (1991) had the lowest estimate of 0.3% which was extrapolated from a cohort of probands less than age fifty with an incidence of 3.1%.

Once the MMR genes were identified and some variants shown to cause hereditary colon cancer, multiple studies completed molecular characterization for incidence estimates. All used similar methodology including MSI testing of all the CRC patient tumours followed by mutation detection only when tumours were microsatellite instable to confirm the diagnosis. Incidence estimates ranged from 0.9-2.7% and included populations from the Mediterranean region, USA and Finland (Salovaara et al., 2000; Aaltonen et al., 1998; Cunningham et al., 2001; Ravnik-Glevac et al., 2000; Percesepe et al., 2001; Samowitz et al., 2001; Goodfellow et al., 2001).

One of the most comprehensive studies, of which the population of this thesis is included, is that published by Woods et al. (2010). This population based study included an incident cohort of 750 CRC patients from the province of Newfoundland and Labrador

(NL). 4.6% were classified as having high familial risk based on fulfillment ACI while 3.6% had germline variants and diagnosed as LS.

1.10 *Genotype-Phenotype Variation*

Genotype includes one's genetic make-up whereas phenotype refers to the actual observable characteristic. More specifically, it includes the micro- and macroscopic expression of an individual's genetic makeup. Many studies have examined correlations between genotype and phenotype for hereditary colorectal cancer. In this instance, the value of understanding this relationship lies in the management and appropriate surveillance of affected individuals. With mutations in either *MLH1* or *MSH2* a classic LS phenotype is expected as MMR function is completely lost. When the other proteins are implicated, however, varying phenotypes exist. For example, an attenuated phenotype is usually associated with the *MSH6* genotype. Endometrial cancer is more frequent but colorectal cancer is less penetrant and it, as well as other LS cancers, tend to occur later in life (Hendriks et al., 2004).

About half of all families clinically defined as LS do not have deleterious variants in the known MMR genes and are MSS. This active research area involves the possibility of identifying new CRC susceptibility genes. A mutation in these novel genes may be associated with a completely different molecular mechanism and, therefore, unable to be detected by MSI and IHC (Zhang, 2008).

To further support this, a study completed by Rovella et al. (2001) demonstrated the existence of familial MSS CRC. One hundred consecutive CRC patients from a clinic

in Italy were evaluated through family history, clinical characteristics, and MSI. Eighteen patients had a positive family history of CRC in a first degree relative. Thirteen of these did not demonstrate MSI, had a later age of onset, and had a higher incidence of second primary tumours relative to non-familial CRCs. No significant difference was found between the frequencies of extraintestinal cancer sites in the MSS familial pedigrees. They concluded that the majority of familial CRCs not attributable to known genes are MSS. As this was not population based, no inferences could be drawn to the general population. The investigators recommended collection and clinical characterization of further families for identification of novel genes.

An interesting study by Park et al. in 2007 looked at the clinicopathological characteristics of colorectal cancer with family history. They found the number of patients with MSI tumours was significantly higher in families affected with colorectal or LS associated cancers than those with other cancers. Seventy percent of the cancers in first and second degree relatives of colon cancer patients were gastric and colon cancer, whereas, only 4% were the other LS associated cancers. This study eliminated families who met the Amsterdam criteria, were FAP, or had vague family histories. The authors concluded that family history of cancer could be used to predict an underlying familial cause.

MSI is the molecular phenotypic expression of germline MMR deficiency. Phenotype differences between MSI and MSS families have also been documented. In the study by Valle et al. (2010), sixty-four Amsterdam I-positive families were included from a Spanish National Cancer Centre. Three generational pedigrees were constructed

from family history, tumour samples were collected and examined for MSI status and MMR protein expression, and MSI families were tested for MMR variants. Twenty-six families (40.6%) demonstrated MSS and thirty eight (59.5%) were MSI. Of the MSI families, 30 out of 38 had an MMR defect identified and 20 of the 26 MSS tumours expressed all proteins. Using family cancer history, MSI families had increased incidence of multiple primary tumours including LS related and other cancers, whereas MSS families had a predominance of only CRC. Additionally, a study by Perea et al. (2010) found MSS tumours were associated with diagnosis at a more advanced age. Higher mortality in patients with MSS tumours suggests MSI tumours have a better prognosis.

The term FCCTX was proposed by Lindor et al. (2005) to describe families who had a clustering of colorectal cancer but whose tumours lacked any DNA MMR gene defect. This term did not define the group as having LS. These families did not share the same cancer incidence as families with LS and relatives in these families had a lower incidence of CRC and the incidence of other cancers in family members did not appear to be increased.

1.11 *Newfoundland Population*

Genetic isolation has decreased considerably in the last century in most European and North American populations. This is due mainly to the development of extensive transportation systems and concentration of the population in urban areas. This trend, however, does not hold for all populations. Some populations still remain genetically

isolated due to a long history of geographical isolation with minimal immigration. Such is the case for the island portion of the Canadian province of NL (Bear et al., 1987).

The Newfoundland population was originally founded in the late 1700s by about 12,000 settlers from localized areas of southwest England and southeast Ireland. Small, isolated communities were first established in Conception Bay and the Southern Shore of the Avalon Peninsula where the local economy was based on the inshore fishery. With the expansion of the population, new fishing communities were established around Newfoundland's natural harbours. The migration pattern followed the coastline further west, on the north and south borders, and eventually up the west and north coasts of the island. The population has continued to expand with a large percentage still residing in these isolated coastal communities with approximately 50% of Newfoundland's population living in communities of fewer than 2,500 inhabitants even now.

Genetically, the Newfoundland population is ideal for studying genetic diseases due to its founder history and geographical isolation. Moreover, Newfoundland families tend to be larger and family records are relatively well detailed allowing family history to be traced back over several generations.

1.12 *Rationale*

The goal of this project was to determine the genetic contribution to colorectal cancer in Newfoundland and Labrador through an analysis of family history and molecular characteristics of CRC in the year 1999.

Just prior to the initiation of this study, a pilot project was implemented examining colorectal cancer cases diagnosed only on the Avalon Peninsula in Newfoundland during two years: 1997 & 1998. Conclusions from this study noted a high proportion of cases with high or intermediate risk family histories (55.1%), and a majority of cases who fulfilled Amsterdam criteria had MSS CRC suggesting genes other than MMR may be involved in hereditary CRC (Woods et al. 2005).

The population base for the current study includes the entire province of NL, rather than the Avalon Peninsula, as in the pilot study, and analyzes a larger sample of incident colorectal cancer cases. As this was a large collaborative study funded by the Canadian Institute of Health Research (CIHR), more resources were available to increase study participation and complete family history data, including cancer verification, and thus provide more accurate and detailed results.

1.13 *Background*

In 2001, a team of investigators in Ontario and NL developed an interdisciplinary health research team (IHRT) aiming to advance knowledge of the determinants, the impact, and the control of colorectal cancer. Six specific research goals were outlined:

1) A case control study to assess whether genes involved in mismatch repair are associated with CRC risk in Ontario and NL.

2) Among the relatives of probands, the mutation-, sex-, and age-specific cumulative risks of CRC in MMR gene mutation carriers will be estimated and compared with

similar risks in Ontario families, and these further compared to risks in families with a Newfoundland founder mutation.

3) In a follow-up study of CRC patients, the prognostic significance of microsatellite instability in colorectal tumours assessed in terms of survival among cases from both provinces.

4) A retrospective study of high risk and low risk individuals, who are relatives of CRC patients, to identify factors that may influence the decision to be screened for CRC by fecal occult blood testing or colonoscopy.

5) The psychosocial and behavioural impact of genetic counselling and testing and of a diagnosis of CRC in a family assessed among high risk and intermediate risk individuals.

6) A genetic risk assessment tool for physicians and an information aid for individuals with a family history of CRC will be developed and evaluated.

(CIHR-IHRT Research Proposal 2000)

For this thesis, data obtained from the incident CRC patients in the first year of data collection from the IHRT study in NL was analyzed. More specifically, an incident cohort study over one year was undertaken utilizing both self reported family history and tumour molecular analyses to determine the genetic contribution to colorectal cancer in NL.

1.13.1 *CRC-IHRT Results*

As outlined in the notes to the reader, results from the five year CRC-IHRT study have been published by Woods et al. and referenced where applicable. An incident cohort of 750 patients with CRC representing 708 different families from the NL population was studied. Eligible patients were those diagnosed under the age of 75 from January 1, 1999 to December 31, 2003. Risk classification was completed as outlined in this study but additional, more detailed colorectal tumour analyses as well as germline mutation screening was also completed. They reported 4.6% of patients fulfilled ACI and 44.6% fulfilled the RBG. Using high, low, and stable tumour microsatellite classifications, MSI-H was demonstrated in 10.7% (78/732). Of these, 42 tumours (53.8%) showed methylation of the *MLH1* promoter, 21 (26.9%) were from patients found to have a deleterious variant in a MMR protein, and the remaining 15 (19.2%) tumours had no cause identified for the MMR deficiency. Twelve different mutations were found in six known CRC-related genes and all mutation carriers demonstrated MSI-H in their CRC tumour. Overall the total proportion of MMR mutations (Lynch syndrome) was 2.7%. Additionally, 61% (17/28) of Amsterdam criteria families did not have a genetic cause identified and 15 fulfilled FCCTX criteria proposing the existence of novel mutations in CRC-predisposing genes. Finally, patients from FCCTX families were found to have a later age of CRC diagnosis and increased occurrence of distal tumour location compared to patients with LS (2010).

1.14 *Objectives*

The main objective of this study was to determine the genetic basis of CRC in NL as defined by family history, molecular pathology of the CRC and molecular genetics in 148 incident cases of CRC aged 20-74 diagnosed in the NL population during one year. In order to achieve this, I estimated the number of instances of possible LS and FCCTX.

The secondary objective was to further elucidate the phenotype of FCCTX compared to possible LS. This was achieved by comparing cancer type incidences between the groups.

Chapter 2: *Materials and Methods*

2.1 *Study Population Selection*

The protocol for the larger study of which this study is a part was designed by researchers in NL and Ontario. Identification of the population for the entire five year study was completed by CRC-IHRT research team members prior to initiation of this thesis following the protocol outlined below. It is important to note, however, that although I did not personally select the population for this study I did gain a thorough understanding of the process through reviewing the detailed protocol provided by the principal investigators, as well as working in conjunction with other team members as a research assistant on this project prior to enrolling as a masters student.

All colorectal cancer patients aged twenty to seventy-four entered into the NL provincial tumor registry, diagnosed from January 1, 1999 through December 31, 2003, were retrieved. Unregistered cases were identified through review of hospital ICD9 codes and through the provincial death statistics. Demographic information including sex, age, address, telephone numbers, as well as attending physicians, and other related cancer histories were extracted from the registry. Corresponding tumour pathology reports of these patients registered with the provincial cancer registry in the province of NL were retrieved and reviewed by expert pathology team members. The following pathological inclusion and exclusion criteria developed by the principal investigators were applied to determine which patients were forwarded to the recruitment team.

2.1.1 *Inclusion Criteria*

Any tumour described as adenocarcinoma (including any mention of signet ring cells but not including carcinoid), adenosquamous, undifferentiated, poorly differentiated, small cell carcinoma, or any combination of the previous from all parts of colon, rectum and anal canal were included. Squamous cell carcinoma was only included if it was definitely arising from the rectum or colonic mucosa. Inclusion of pseudomyxoma was only included if accompanied by adenocarcinoma.

2.1.2 *Exclusion Criteria*

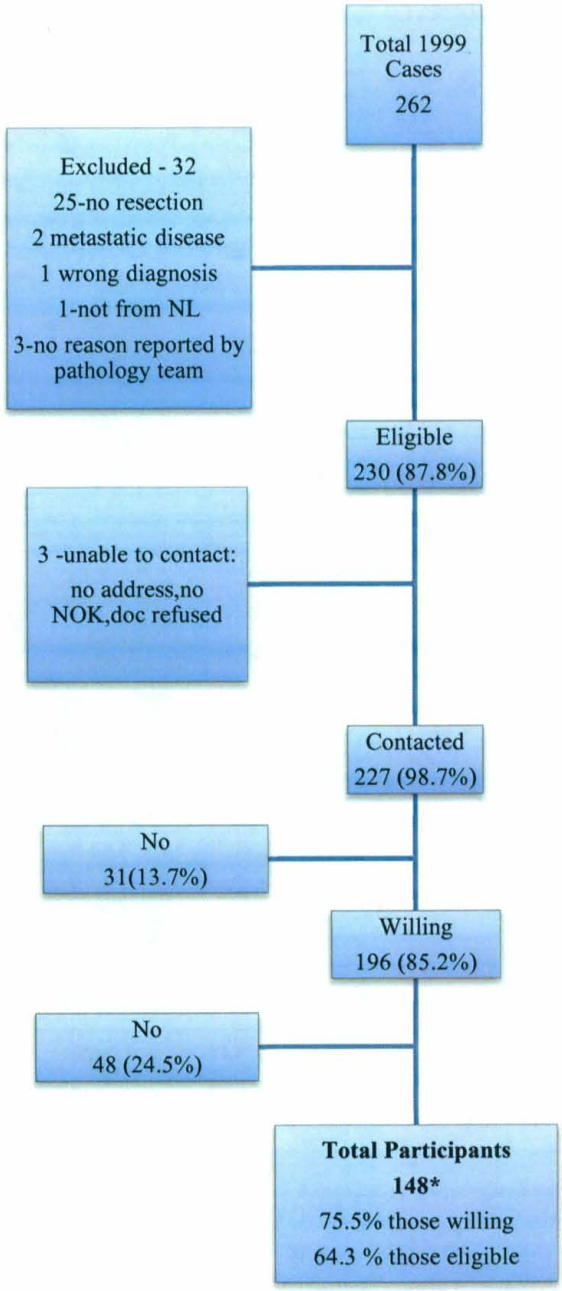
Any tumour that was squamous cell carcinoma of the anal canal, lower rectum or lacked certainty if it originated from colonic or rectal mucosa were excluded. Additionally, all usual carcinoid tumours including goblet cell carcinoid and all pseudomyxoma that were not accompanied by adenocarcinoma were excluded.

2.1.3 *Proxy Identification*

For individuals who were deceased, a proxy (next of kin) was identified and contacted, by the recruitment team, to act on the patient's behalf. Names were available through the provincial cancer registry, the physician's office, or through a follow up phone call to the proband's family. First contact with the proxy was made through the proband's physician unless the proxy was identified during a follow up call; at which time verbal consent was obtained.

2.1.4 *Recruitment*

Two hundred and sixty-two cases of CRC were registered in the province of NL during the year 1999 with one hundred forty-eight having completed family history questionnaires (FHQ) required for this analysis. Figure 1 summarizes the recruitment process. All two hundred sixty-two pathology reports were obtained through the provincial cancer registry. A total of thirty two were excluded: twenty-five were excluded because there was no resection of the tumour, two because the occurrence was the result of metastatic disease, one was due to a wrong diagnosis, one was not a NL patient, and for three cases no reason was noted in the database by the pathology team member. Of the remaining two hundred thirty eligible participants contact could not be made with three patients; no current address or contact information could be found for one, another was deceased and no doctor or next of kin identified, and in the third instance the doctor said the patient was not suitable for the study and refused to send the contact letter. One hundred ninety-six patients (83.3%), of the two hundred twenty-seven patients contacted, were initially willing and FHQs were mailed. Forty-eight (24.5%) patients declined participation after receiving the FHQ while the remaining one hundred forty eight questionnaires were returned (75.5%). Overall, for the first year of data, 98.7% of those eligible were contacted and 64.3% of those eligible completed the FHQ. After collection of the information required for this study, eight probands chose to withdraw from the CRC-IHRT study. Permission was granted from all of these individuals to use any already collected data including molecular analysis of the tumour blocks. He/she did not wish to be contacted in the future to complete any further aspects



* eight withdrew from the study but permission was obtained for use of their data.

Figure 1: Summary of Recruitment of Study Participants

of the CRC-IHRT study. Consequently, 75.5% of those willing and 64.3% of those eligible are included in this study.

2.2 *Data Collection*

I participated in all aspects of data collection however my research was completed as a member of the recruitment team and was not restricted to only participants diagnosed in 1999.

Initial contact with participants was made through the proband's physician. More specifically, each proband identified by the provincial cancer registry had a list of physicians associated with him/her. After participation agreement from the acting physician, a patient-addressed letter and reply card were forwarded by the recruitment team to be signed and mailed on the study's behalf. The self-addressed reply card was returned by the physician to indicate the invitation letter had been mailed. Two weeks post mail-out; the patient was contacted by a registry team member and asked whether he/she was interested in participating. If willing, the recruitment team mailed the participant a package which included the consent forms, details of the study as well as the FHQ. Once consent was obtained, participants completed the FHQ, which was used to assess familial risk, and samples of his/her tumor block were collected by the molecular team to undergo molecular analysis including microsatellite instability testing and immunohistochemical staining.

Demographic information including sex, mean age of diagnosis, CRC tumor site, and residence for non-participants was also collected from the registry which I compiled into a working data set of all probands diagnosed with CRC during the year 1999. I

compared this information to that of participants to determine the representativeness of my sample.

2.2.1 *Family History Questionnaire*

This questionnaire was designed so a three-generation pedigree could be constructed. More specifically, questions were asked regarding the family cancer history of grandparents, parents, siblings, and children. In addition, names of other family members who had been diagnosed with cancer and age of cancer diagnoses were requested. Following this mail out, regular follow up took place via phone to increase the response rate. A phone or in-person interview was available and a toll free phone line was accessible to all participants if assistance was required. Upon receipt of the FHQ, a family pedigree was constructed using the computer package Cyrillic 3.

2.2.2 *Verification of Cancers*

This portion of the methods was completed in partnership with the genetic counselor Angie Batstone.

In order to ensure accuracy of the history provided and subsequent risk assessment, verification of all hereditary associated cancers in relatives was attempted in first and second degree relatives. The cancers included were colon, endometrial, liver, stomach, brain, and ovary. Any non-associated and unknown cancers were only attempted to be verified if they directly affected risk classification. For those relatives who were alive, a family address sheet was mailed to the proband asking him/her to contact relatives requesting permission for the research team to contact the relative. For

those who were deceased and for those whom permission was obtained, a request for information (ROI) form was mailed which required signature of the living relative or from the next of kin of those who were deceased. Upon receipt of the signed ROI, relevant medical records were requested. Two hundred seventy-seven reported cancers were attempted to be verified, of which 44% were successful (see section 3.2.5 for detailed results). Some families had been previously identified by the Provincial Medical Genetics Program (PGMP), which is responsible for all the regional genetic screening. This data was accessed, with permission, where applicable for any cancers that were previously confirmed. Verification was often not possible due to inability to obtain records of cancers diagnosed outside the province or those diagnosed before records at the corresponding hospitals were retrievable. In addition, some probands did not have any contact with the next of kin, or relatives declined participation preventing cancer verification.

2.2.3 *Health Regions*

At the time of this study, NL had six distinct health regions: St. John's, Eastern, Central, Western, Grenfell and Labrador. The boundary of each region was defined by the Department of Health and illustrated in Figure 2. Individuals from Grenfell and Labrador health regions were combined into one group. Geographically these two regions are adjacent and while they may be the largest in size they represent the fewest number of participants. Each proband was assigned to the health region corresponding to the personal address provided by the provincial tumor registry which was recorded at the time of registration.

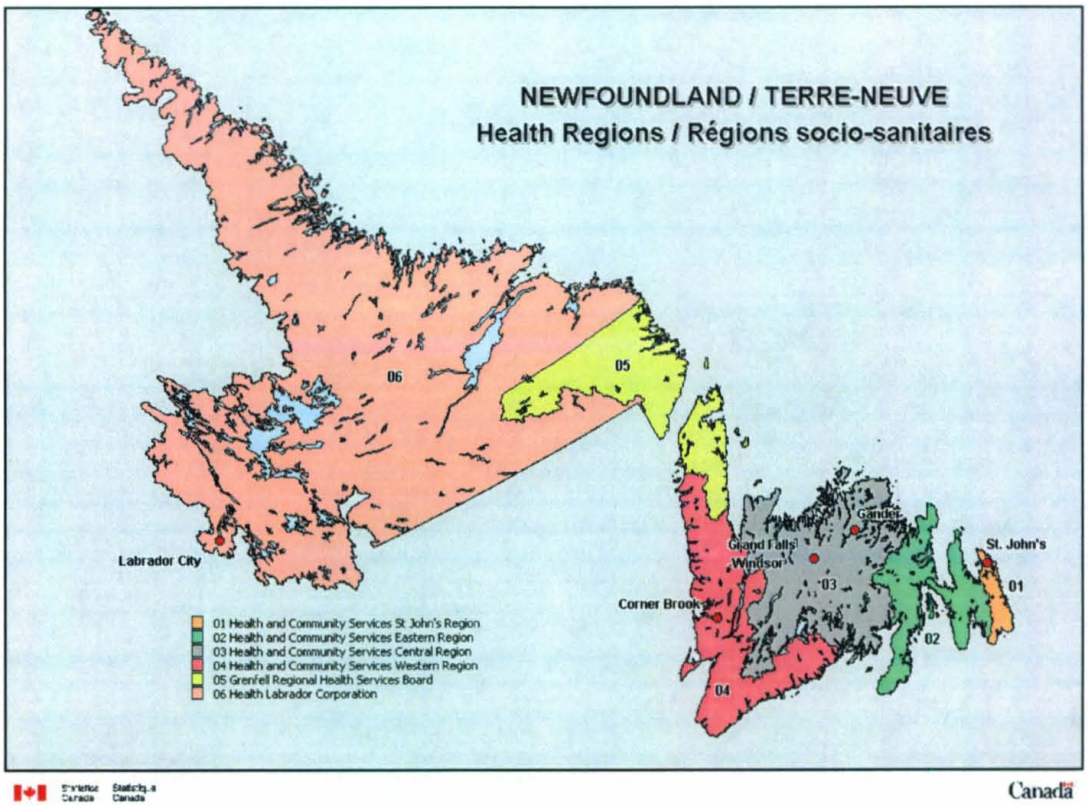


Figure 2: Six Defined Health Regions of Newfoundland and Labrador

http://www.statcan.gc.ca/concepts/health-sante/maps-cartes/nfld-eng.jpg

2.3 *Family History Risk Assignment*

A three generation pedigree was constructed for each participant according to reported cancers, and risk assignment (the likelihood he/she might be carrying an inherited deleterious variant) was conducted. I determined familial risk classification in accordance with internationally recognized criteria as outlined by the CRC-IHRT study. In partnership with Dr. Jane Green, all pedigrees were reviewed for accuracy of risk classification and any changes to risk classification following family history updates were completed under the direction of Dr. Jane Green.

2.3.1 *High Risk*

Family history high risk families were classified according to ICG-HNPCC established criteria for ACI and ACII (see 1.2.1 and 1.2.2). A third classification of high risk was defined by altering the age criteria and tumor type for ACII. This was classified as age and cancer-modified Amsterdam criteria (ACMAC) and is defined by a family meeting all Amsterdam II criteria with the age limit raised from 50 to 60 and the extra-colonic tumor list expanded to include all the tumors of the RBG (see 1.2.4). ACMAC incorporated high risk families that had been excluded based on the arbitrary age and tumour list of the Amsterdam criteria but clearly demonstrated characteristics of autosomal dominant disease; every affected individual has an affected biological parent, there is no skipping of generations, and males and females have an equally likely chance of inheriting the mutant allele and being affected.

2.3.2 *Intermediate Risk*

Intermediate risk (IR) was assigned to families who did not fulfill ACI or II but did fulfill the clinical criteria of the RBG. The use of the RBG to confer an intermediate familial risk classification was part of the design of the CRC-IHRT study. While they are the third clinical criteria to identify possible LS families, the RBG are the least stringent to identify those with germline variants in one of the MMR genes. They were originally designed to identify individuals whose colorectal tumours should be tested for MSI, not specifically identify families that meet clinical criteria for LS (see section 1.2.4 for specific criteria). Additionally, for the purposes of this study, criterion 3 (MSI-histology of tumours) was not utilized. This criterion is based on pathology analysis that is not routinely assessed or reported in practice. Those that satisfied more than one criterion were also classified according to the criteria most suggestive of a Mendelian inheritance pattern. Specifically, compared to the other criteria, revised Bethesda criteria 4 and 5 have the highest family frequency of LS associated or CRC cancers represented in successive generations thereby most suggestive of a Mendelian inheritance pattern. Therefore, revised Bethesda 4 or 5 criteria was the sole classification when these criteria were met in addition to others.

2.3.3 *Low Risk (Sporadic)*

All remaining probands who did not meet either high or intermediate risk criteria were categorized as low risk. These families were further classified as informative or non-informative according to the reported family history based on the definition outlined

in the similar study completed for the Avalon Peninsula. Informative families were defined as those with six or more family members in two successive generations potentially at 50% risk of inheriting CRC who had survived until at least age 60. As LS follows a Mendelian inheritance pattern and age of onset is usually before sixty, this definition ensures a pedigree has enough information to make an accurate risk classification based on the information available (Curtis, 2003).

2.4 *Genetic Counseling*

As this area requires specialized training I shadowed and assisted the genetics nurse to learn the procedure. Letters to the patients included in this study were reviewed as well as the necessary screening protocols.

All participants were offered genetic counseling. Low risk patients were mailed a letter outlining their risk and recommending case specific screening strategies as well as offering counseling. High and intermediate risk patients were invited directly for counseling and then mailed a letter which outlined their risk. Each chart was reviewed by genetics nurse Angie Batstone and Dr. Jane Green.

2.5 *Tumor Blocks*

Tumours in paraffin blocks were requested by the pathology team for all participants after we advised consent had been received. Hospitals where the patient had undergone surgery were contacted with the corresponding tumor identification code. They forwarded a sample of the tumor and normal tissue. Regular consultations occurred with the pathology team regarding tumours received and those outstanding.

2.6 *Molecular Analysis*

All tumours were analyzed for microsatellite instability, by Dr. Roger Green's lab, and protein expression using immunohistochemistry staining, by Dr. Des Robb's lab. Information for this study was received after being compiled by Dr. Roger Green in a confidential excel spreadsheet.

2.6.1 *Microsatellite Instability*

Five microsatellite markers were used to classify all tumours as microsatellite stable or unstable by comparing matched normal and tumor DNA. The markers used were BAT25, BAT26, D2S123, D5S346, D17S250. In this study, tumours were classified as microsatellite stable if fewer than two markers demonstrated instability or microsatellite high if two or more markers were unstable. MSI-low has been previously used when one marker was found to be unstable however, as outlined in section 1.5, no convincing molecular or physiologic difference has been found between MSI-low and MSS tumours and therefore MSI-high is referred to as simply MSI and MSI-low as MSS.

2.6.2 *Immunohistochemistry Staining*

Immunohistochemical staining was performed on each tumor for three mismatch repair proteins; *MLH1*, *MSH2*, *MSH6*. Testing for *PMS2* was not completed for this study. If protein expression was positive, the tumor was classified as having no gene dysfunction for that particular protein. If the tumor did not exhibit protein expression the tumor was classified as having gene dysfunction for that particular protein.

2.7 *Genetic Classification*

Using IHC results, I further classified high risk families as probable LS, FCCTX or age modified FCCTX. Families of patients with tumours demonstrating MMR protein deficiency were classified as possible LS, while those who fulfilled AC and had no MMR deficiency were referred to as FCCTX. Finally, families of patients whose tumours expressed all proteins and fulfilled the ACMAC high risk criteria were considered age modified-FCCTX.

Families can only be definitely classified as LS when a deleterious variant is found in a MMR gene. While this analysis was being completed by another molecular team only MSI and IHC results were designed to be included in this study as the basis for genetic classifications. Although the detailed mutational analysis is beyond the scope of this thesis, results from the five year study in its entirety have since been published by Woods et al. (2010) and reference has been made where applicable. Access to the detailed results was not part of the data collection for this thesis.

2.8 *Analysis*

Molecular information and family history were compiled and analyzed using the Statistical Package for Social Sciences (SPSS). Comparison of willing participants to those that declined was completed using chi-square to compare to sex, CRC tumour site and residence and independent samples t-test to compare mean age of diagnosis. Chi-square analysis was also used to test if there was any difference between health regions in the frequency of low, intermediate or high risk families. No more than 20% of the cells

had an expected frequency less than 5, meeting the assumption for the chi-square test of independence. Significance was set at $p<0.05$.

Clinical cancer phenotype was also compared between possible LS and FCCTX families. First and second degree relatives were included to compare age of diagnosis and frequency of occurrence of CRC and associated hereditary cancers. Age at diagnosis for those with cancer occurrence or age of last follow up for those without cancer was used for analysis. Hereditary cancers were those defined by the RBG including colorectal, endometrial, stomach, ovarian, pancreas, ureter, renal pelvis, biliary tract, brain (glioblastoma), small intestine, sebaceous gland adenomas and carcinomas and keratocanthomas. Kaplan-Meier time-to-event survival curves were constructed to demonstrate differences in age of cancer onset between possible LS and FCCTX families. Statistical significance was determined using Mantel-Cox and Breslow with $p<0.05$.

2.9 *Ethical Considerations*

Approval for the larger CRC-IHRT study was obtained from the Human Investigations Committee at Memorial University of Newfoundland. All identified individuals were entered into a restricted, password-protected database and all corresponding files kept in a locked filing cabinet. All individuals were initially contacted by their physician and informed consent was obtained only after he/she demonstrated an interest in participation. Participants were free to withdraw at any time or to decline to complete any portion of the study. Completing requested questionnaires and/or contacting requested family members may have been time consuming or difficult

for the patient and therefore was completed at the participant's convenience with or without assistance. To maintain patient confidentiality all molecular analyses were completed using only corresponding patient study identification numbers.

Chapter 3: *Results*

3.1 *Characteristics of Participants*

Of the one hundred forty-eight willing participants providing complete information, ninety (60.8%) were living and fifty-eight (39.2%) were deceased and represented by proxies. There were sixty-one (41.2%) females and eighty-seven (58.8%) males. The age range of CRC patients was 20-74 years with a mean of 61 years of age. The total population at risk in 1999, aged 20-74, was five hundred forty thousand, eight hundred ninety-five (540,895) for an incidence of colon cancer of 0.0484%. Distribution of incidence according to the five defined health regions of NL revealed forty-six (31.1%) participants were from the St. John's region, twenty-nine (19.6%) from the Eastern region (exclusive of St.John's), forty-three (29.1%) from Central, nineteen (12.8%) from Western, and eleven (7.4%) from the Grenfell/Labrador region. The tumour site for one hundred (67.6%) of the participants was the proximal colon and for forty-eight (32.4%) the distal colon including the recto-sigmoid, rectum, and anus.

Table 1 summarizes the characteristics of participants and non-participants. No significant differences between participants and non-participants were found when sex, mean age of diagnosis, CRC tumour site, and residence were compared, suggesting the study population for this thesis is representative of the whole population.

Table 1: Characteristics of Participants (n=148) and Non-Participants (n=82).
Statistical comparison of participants to non-participants using chi-square and independent samples t-test to test for recruitment bias.

	Participants	**Non-Participants**
Alive	90 (60.8%)	44 (53.7%) *
Deceased	58 (39.2%)	38 (46.3%) *
Male	87 (58.8%)	48 (58.5%) *
Female	61 (41.2%)	34 (41.5%) *
Mean Age CRC diagnosis	61	62.6 **
Age Range CRC diagnosis	20-74	22-74 **
Tumour Site		
Colon	100 (67.6%)	53 (64.6%) *
Rectum, Recto-sigmoid, Anus	48 (32.4%)	29 (35.4%) *
Place of Residence		
St. John's	46 (31.1%)	25 (30.5%) *
Eastern	29 (19.6%)	14 (17.1%) *
Central	42 (29.1%)	24 (29.3%) *
Western	19 (12.8%)	15 (18.3%) *
Grenfell/Labrador	8 (7.4%)	4 (4.8%) *

*p>0.05 between groups, Chi-Square Analysis

**p>0.05 between groups, Independent Samples T-test

3.2 *Family History Study*

3.2.1 *Risk Classification*

A three generation pedigree was constructed for each participant and familial/clinical risk was assigned according to family reported cancers. Nine probands were found to be relatives of one another and required unique risk consideration: four were first degree relatives (siblings), two were second degree relatives (uncle and nephew), two were fourth degree relatives, and one pair of the siblings also had a third degree relative who was a proband. The first, second, and third degree relatives were considered to have the same familial risk and therefore were only entered once in the family history analysis. The fourth degree relatives, however, had different maternal versus paternal sides at risk and it was determined by Dr. Jane Green that these probands required separate risk classifications and therefore were independently counted. Consequently, while there were one hundred and forty eight probands, there were a total of one hundred and forty-four families reviewed for family history risk classification (Table 2).

Eighteen families (12.5%) were classified as high risk for possible LS. Four (22.2%) of these fulfilled ACI and fourteen (77.8%) fulfilled ACMAC.

Forty-eight (33.3%) families were classified as intermediate risk according to the RBG, exclusive of the eighteen high risk families. As previously indicated, Bethesda criterion 3 (colorectal cancer with MSI-H histology diagnosed in a patient < 60 years of age) was not utilized since risk classification in this study was according to family history

Table 2: Frequencies of Family History Classifications Organized by Risk Category.

	Families	%
High Risk	**18**	**12.5**
Amsterdam I	4	2.8
Amsterdam II	0	0
Age/Cancer Modified Amsterdam	14	9.7
Intermediate Risk	**48**	**33.3**
Bethesda 1	10	6.9
Bethesda 2	5	3.5
Bethesda 4	3	2.1
Bethesda 5	16	11.1
Bethesda 1&5	5	3.5
Bethesda 1&2&5	1	0.7
Bethesda 2&5	6	4.2
Bethesda 4&5	1	0.7
Bethesda 2&4	1	0.7
Low	**77**	**53.5**
FAP	**1**	**0.7**
TOTAL	**144**	**100**

only. Sixteen (33.3%) probands were IR based on having two first or second degree relatives with RBG defined tumours (Bethesda criterion 5, B-5). Three (6.3%) probands met Bethesda criterion four (B-4) based on having a first degree relative with a RBG-defined cancer less than 50 years of age. Five (10.4%) fulfilled Bethesda criterion two (B-2) based on the presence of synchronous or metachronous RBG defined cancers, and ten (20.8%) fulfilled the criterion of a CRC tumour in a proband less than 50 years of age (Bethesda criterion 1, B-1). Fourteen probands (29.2%) classified as IR fulfilled multiple Bethesda criteria. Six met B-2 and B-5, one met B-4 and B-5, one met B-2 and B-4, and one met B-1, B-2, and B-5. Final classification was according to the criteria most suggestive of a Mendelian inheritance pattern, specifically using the criteria with the highest frequency of RBG defined or CRC cancers in successive generations. Any proband that satisfied Bethesda criteria 5 in addition to others received Bethesda 5 as his/her sole risk and the one proband satisfying Bethesda 2 and 4 was classified as Bethesda 4 (Table 3). Overall, twenty nine (60.4%) families met B-5, four (8.3%) met B-4, five (10.4%) met B-2 and ten (20.8%) met B-1.

Seventy seven (53.5%) families did not meet high or intermediate risk and were classified as low risk.

One family (0.7%) was a known FAP family as identified by Dr. Jane Green. FAP is a known hereditary CRC syndrome associated with the *APC* gene and therefore was not risk classified as high, intermediate or low and is represented as a family history classification only.

Table 3: Final Family History Classification Frequencies. Intermediate risk families fulfilling more than one Bethesda criteria were classified according to the criteria most suggestive of a Mendelian inheritance pattern.

	Families	**%**
High Risk	**18**	**12.5**
Amsterdam I	4	2.8
Amsterdam II	0	0
Age/Cancer Modified Amsterdam	14	9.7
Intermediate Risk	**48**	**33.3**
Bethesda 1	10	6.9
Bethesda 2	5	3.5
Bethesda 4	4	2.8
Bethesda 5	29	20.1
Low	**77**	**53.5**
FAP	**1**	**0.7**
TOTAL	**144**	**100**

3.2.2 *Low Risk Family Informativeness*

Fifty-two low risk families (67.5%) were classified as informative and twenty-five (32.5%) uninformative. Of the twenty-five uninformative families, twelve had fewer than six family members in two successive generations potentially at 50% risk of inheriting any autosomal dominant cancer, five had fewer than six family members at risk that had survived to sixty years of age or developed cancer, and eight had insufficient family history.

3.2.3 *Distribution of Family Risk According to Provincial Regions*

Each participant was assigned to one of the five health regions in Newfoundland and Labrador according to the proband's address provided at the time of inclusion into the study. St. John's, Eastern, Central, Western, and Grenfell/Labrador health regions were used whose boundaries follow those outlined by the Department of Health. Table 4 summarizes incidence rates of colon cancer in 1999 according to health region for ages 20-74; forty-six cases (31.1%) were found in the St. John's region, twenty-nine (19.6%) in the Eastern, forty-two (29.1%) in Central, nineteen (12.8%) in Western, and eleven (7.4%) in the Grenfell/Labrador region. Regional incidences per 100,000 were calculated and found to be 36.0, 36.4, 58.9, 31.9, and 8.6 respectively. Using chi-square, there was no significant difference between regional incidence rates with $p = 0.406$. Further characterization using family history risk classification (high, intermediate and low risk) according to health region was also completed and is summarized in Table 4A. Of the total number of colon cancers per health region, Western had the highest frequency

Table 4: Incidence Rates of Colon Cancer in 1999 According to Provincial Region for ages 20-74. Statistical comparison of incidence rates of colon cancer according to NL provincial health regions for ages 20-74 using chi-square.

	St. John's	Eastern	Central	Western	Grenfell/ Labrador
Total Colon Cancers	**46(31.1%)**	**29(19.6%)**	**42(29.1%)**	**19(12.8%)**	**11(7.4%)**
Population At Risk age 20-74	**127,682**	**79,753**	**72,980**	**59,644**	**28,492**
Incidence Rate*	36.0/100,000	36.4/100,000	58.9/100,000	31.9/100,000	8.6/100,000

*p=0.406

Source: Statistics Canada, Population Estimates, 1999. Prepared by the Research and Evaluation Department at the Newfoundland and Labrador Centre for Health Information.

Table 4A: Comparison of Frequencies of Family History Risk Classifications Between Provincial Health Regions. Statistical comparison of high, intermediate and low familial risk frequencies between NL health regions using chi-square. Risk frequencies are calculated per region using the total number of regional colon cancers as the denominator.

	St. John's	Eastern	Central	Western	Grenfell/ Labrador	
Total CRC	**46**	**29**	**42**	**19**	**11**	**p-value**
High	**6(13.0%)**	**4(3.8%)**	**3(7.1%)**	**4(21.1%)**	**1(9.1%)**	**0.999**
Amsterdam I	2	0	1	1	0	
ACMAC	4	4	2	3	1	
Intermediate	**13(28.2%)**	**9(31.0%)**	**16(38.1%)**	**8(42.1%)**	**6(54.5%)**	**0.406**
Low	**27(58.7%)**	**16(55.2%)**	**23(54.8%)**	**7(36.8%)**	**4(36.4%)**	**0.406**

(21.1%) of high risk family history CRC; while St. John's had the highest frequency (58.7%) of low risk family history classified CRC. Although no statistically significant differences between incidence were found ($p>0.05$) when the frequencies of high, intermediate, and low family history risk classifications were compared between the five health regions, clinically these differences may be important.

3.2.4 *Summary of Family History Reported Cancers*

Reported cancers in first and second degree relatives of the one hundred forty-eight probands are summarized, according to assigned family history risk, in Tables 5 to 8. CRC was the most frequently occurring tumour in the high and intermediate risk families while breast and lung were most common in the low risk.

Of seventy-seven relatives in the ACI families, twenty-five (32.5%) were reported to have developed cancer. The mean age for occurrence was 45.7 years with a range of 7-73 years of age (three relatives had unknown age of occurrence). A total of eight different cancer types and one unknown were reported with the most common extra-colonic tumour being endometrial. Nine relatives were reported as having two cancers and were counted separately in the frequency table. Therefore, a total of thirty-four cancers were reported (Table 5).

The ACMAC family history risk group had a total of two hundred sixty-eight relatives. Sixty-eight (25.4%) were reported as developing cancer with the mean age of occurrence being 59.8 years and the range 11-89 years of age (nine relatives had an unknown age of occurrence). Twenty-two types of cancer were reported with stomach

Table 5: Frequency of Family History Reported Cancers for Amsterdam Criteria I Families According to Specific Cancer Type Categories (n = 4).

Cancer Type	# reported	% of reported
CRC	**15**	**44.1**
AMSTERDAM II	**8**	**23.5**
Endometrial	6	
small bowel	2	
Ureter	0	
renal pelvis	0	
BETHESDA	**6**	**17.6**
Stomach	4	
ovary	2	
Pancreas	0	
biliary tract (liver,gallbladder,bile duct)	0	
brain (glioblastoma)	0	
sebaceous gland	0	
OTHER	**4**	**11.8**
Skin	2	
Breast	1	
Prostate	1	
UNKNOWN	**1**	**2.9**
TOTAL	**34***	**100**

*Nine relatives were reported as having two primary cancers:

1 CRC and ovary	2 CRC and stomach
2 CRC and small bowel	1 endometrial and breast
2 CRC and endometrial	1 skin and prostate

Table 6: Frequency of Family History Reported Cancers for Age and Cancer Modified Amsterdam Criteria Families According to Specific Cancer Type Categories (n = 14).

Cancer Type	# reported	% of reported
CRC	**28**	**38.4**
AMSTERDAM II	**1**	**1.4**
Endometrial	1	
small bowel	0	
Ureter	0	
renal pelvis	0	
BETHESDA	**13**	**17.8**
Stomach	7	
Ovary	3	
Pancreas	1	
biliary tract (liver,gallbladder,bile duct)	2	
brain (glioblastoma)	0	
sebaceous gland	0	
OTHER	**25**	**34.2**
Skin	4	
Lung	2	
Breast	2	
Cervix	1	
Prostate	6	
Kidney	2	
Leukemia	1	
Bone	1	
Lymphoma	4	
Thyroid	1	
Neuroblastoma	1	
UNKNOWN	**6**	**8.2**
TOTAL	**73***	**100**

*Four relatives were reported as having multiple primaries:

1 CRC and ovary

1 CRC and kidney

1 CRC, stomach, and prostate

1 CRC and prostate

Table 7: Frequency of Family History Reported Cancers for Intermediate Risk Families According to Specific Cancer Type Categories (n = 48).

Cancer Type	# reported	% of reported
CRC	**53**	**31.9**
AMSTERDAM II	**5**	**3.6**
Endometrial	3	
small bowel	1	
Ureter	1	
renal pelvis	0	
BETHESDA	32	18.8
Stomach	21	
Ovary	2	
Pancreas	4	
biliary tract (liver,gallbladder,bile duct)	5	
brain (glioblastoma)	0	
sebaceous gland	0	
OTHER	**80**	**39.1**
Skin	10	
Lung	21	
Breast	18	
Cervix	4	
Prostate	6	
Kidney	3	
Leukemia	1	
Bone	2	
Lymphoma	3	
Brain	2	
tongue/mouth/throat	6	
Testicular	1	
Bladder	1	
Myeloma	2	
UNKNOWN	**9**	**6.5**
TOTAL	**174***	**100**

*eleven relatives reported with multiple cancers:

1 kidney and lung	1 stomach and lung	1 stomach and breast	2 liver and lung
3 CRC and lung	1 breast and bladder	1 stomach and skin	1 CRC, stomach and lung

Table 8: Frequency of Family History Reported Cancers for Low Risk Families According to Specific Cancer Type Categories (n = 77).

Cancer Type	# reported	% of reported
CRC	**19**	**9.9**
AMSTERDAM II	**4**	**2.1**
Endometrial	4	
small bowel	0	
Ureter	0	
renal pelvis	0	
BETHESDA	**25**	**13**
Stomach	15	
Ovary	1	
Pancreas	2	
biliary tract (liver,gallbladder,bile duct)	7	
OTHER	**122**	**63.5**
Skin	14	
Lung	24	
Breast	24	
Cervix	6	
Prostate	12	
Kidney	5	
Leukemia	5	
Bone	4	
Lymphoma	6	
Thyroid	2	
Brain	1	
Bladder	6	
Myeloma	2	
tongue/throat	**7**	
Testicular	1	
Esophagus	1	
vocal cords	1	
pseudomyxoma peritonei	1**	
UNKNOWN	**22**	**11.5**
TOTAL	**192***	**100**

*five relatives were reported with multiple cancers:

1 stomach and vocal cords | 1 kidney and lung | 1 breast and bone

1 liver, pancreas, stomach | 1 stomach and prostate

***primary was confirmed as unspecified but stomach excluded

being the most common extra-colonic cancer. Four relatives were reported as having multiple cancers to make a total of seventy-three reported cancers (Table 6).

Eight hundred sixty relatives were included for the IR families. Of these, one hundred seventy, (20%), were reported as developing cancer. The mean age of diagnosis was 60.7 years with a range of 22-90 years of age. Age of diagnosis was missing for forty-one relatives. Twenty-two different types of cancer were reported with the most common extra-colonic occurrence being stomach. Eleven relatives had multiple cancers to make a total of one hundred eighty-two reported cancers (Table 7).

Low risk families had a total of one thousand two hundred eighty-three relatives included. One hundred eighty-five (14.4%) relatives were reported as developing cancer with a mean age of 63.4 years for occurrence and a range of 18-91 years of age (age of diagnosis was missing for forty relatives). Twenty-seven different types of cancer were reported and the most common extra-colonic cancer was lung. Multiple cancer occurrences (one with three cancers and five with two) were reported for six relatives making a total of one hundred ninety-two reported cancers (Table 8).

3.2.5 *Confirmation of Family Reported Cancers*

All reported cancers were reviewed with the genetic counsellor and confirmation attempted only for those directly linked to the patient's family history risk classification, such as colorectal or other ACII cancers. Of the four hundred forty-seven first and second degree relatives reported as having cancer in this study, confirmation was attempted for two hundred seventy-seven (62.0%). Confirmation was not required for

one hundred seventy relatives (38%). One hundred and twenty-three (44.4%) of the attempts were successful, leaving one hundred fifty-four (55.6%) unable to be confirmed. Ninety-eight of those confirmed were through direct contact with the patient, or next of kin for those deceased. An additional twenty-five were confirmed through the PMGP using previously identified families. The main reasons for lack of confirmation for the 55.6% failed attempts included, no records being available (67), the requested ROI forms were not completed (32), and refusal to give permission to review medical records (30). Other reasons included: patient withdrew (6), records not obtainable from the hospital (10), and no next of kin available (9). See Table 9 for a summary of these results.

Confirmation status was also organized according to family history risk classification and is summarized in Table 10. Of the ninety-three relatives reported with cancer for high risk families, forty-eight (51.6%) were confirmed, twenty-nine (31.2%) not confirmed and sixteen (17.2%) did not require confirmation. One hundred sixty-nine IR family members with cancer were identified of which forty-nine (29.0%) were confirmed, eighty (47.3%) unable to be confirmed, and forty (23.7%) cancers not requiring confirmation. One hundred eighty-five low risk relatives with cancer were reported with confirmation for twenty-six (14.1%) of the seventy-one cancers attempted while one hundred fourteen (61.6%) did not require confirmation. As previously indicated, we only attempted cancer confirmation for relatives that directly impacted risk classification.

Table 9: Summary of Cancer Confirmation for Family History Reported Cancers for First and Second Degree Relatives.

	N	% of TOTAL
No Confirmation Attempt	**170**	**38**
Confirmed	**123**	**27.5**
Proband	98	
PMGP	25	
Not Confirmed	**154**	**34.5**
No records	67	
Not completed	32	
Permission not given	30	
No next of kin	9	
Can't obtain from hospital	10	
Patient withdrew	6	
TOTAL	**447**	**100**

Table 10: Summary of Cancer Confirmation for Family History Reported Cancers for First and Second Degree Relatives According to Family History Risk Classification.

	High	IR	Low
Confirmed	48 (51.6%)	49 (29.0%)	26 (14.1%)
Not Confirmed	29 (31.2%)	80 (47.3%)	45 (24.3%)
No Confirmation Attempt	16 (17.2%)	40 (23.7%)	114 (61.6%)
TOTAL	**93**	**169**	**185**

3.3 *Molecular Laboratory Study*

3.3.1 *Tumour Collection*

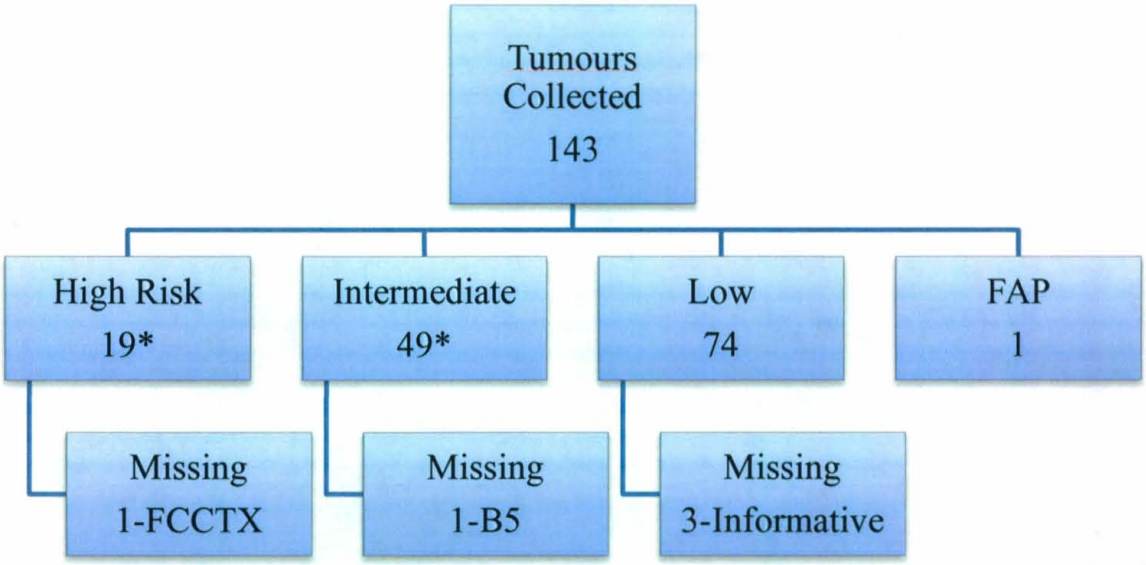
One hundred forty-three tumours were collected for one hundred forty-eight probands/one hundred forty-four families. As only five tumours were unable to be collected, 96.6% of probands representing one hundred forty families (97.7%) had tumours for molecular analysis. Of the five unattainable tumours, one was for a proband in the age modified FCCTX group, one for an IR proband (Bethesda criterion 5) and three for probands in the informative low risk group (Figure 3).

3.3.2 *Microsatellite Analysis*

Analysis was complete for one hundred forty-one (98.6%) tumours and incomplete for two (1.4%); one from a proband in the age modified FCCTX group and one for an IR proband. Table 11 summarizes the results. One hundred twenty-six (88.1%) tumours were microsatellite stable (MSS) and fifteen (10.6%) had microsatellite instability.

3.3.3 *Immunohistochemistry Results*

Immunohistochemistry analysis was completed for one hundred forty-three (100%) tumours and is summarized in Table 12. One hundred twenty-four tumours (85%) had positive expression for all three of the proteins. Thirteen cases (9.1%) had negative expression for one protein and positive expression of the remaining proteins; eleven (7.7%) were *MLH1* negative including one having inconclusive staining for



* 2 probands from the same family

Figure 3: Summary of Tumour Collection According to Family History Classification

Table 11: Summary of Microsatellite Analysis Results for All Tumours Collected.

	N	%	% Total
Complete			
MSS	126	89.4	88.1
MSI-H	15	10.6	10.5
Total	**141**	**100**	**98.6**
Incomplete	**2**		**1.4**
TOTAL	**143**		**100**

Table 12: Summary of Immunohistochemistry Analysis Results for All Tumours.

	N	%
Full Protein Expression	**124**	**86.7**
Protein Absence		
MLH1	11*	7.7
MSH2	1	0.7
MSH6	1	0.7
Total	**13**	**9.1**
Inconclusive Staining with remaining proteins having positive expression		
MLH1	1	0.7
MSH6	4	2.8
MLH1 & MSH6	1	0.7
Total	**6**	**4.2**
TOTAL	**143**	**100**

* one tumour had inconclusive results for MSH6

MSH6, one (0.7%) deficient for *MSH2*, and one (0.7%) deficient for *MSH6*. A total of six tumours had inconclusive results. Five tumours were inconclusive for one protein; one *MLH1* and four *MSH6* with the other proteins expressed. One tumour was inconclusive for two proteins; *MLH1* and *MSH6* while expressing *MSH2*.

3.4 *Family History Risk Classification and Molecular Study Correlation*

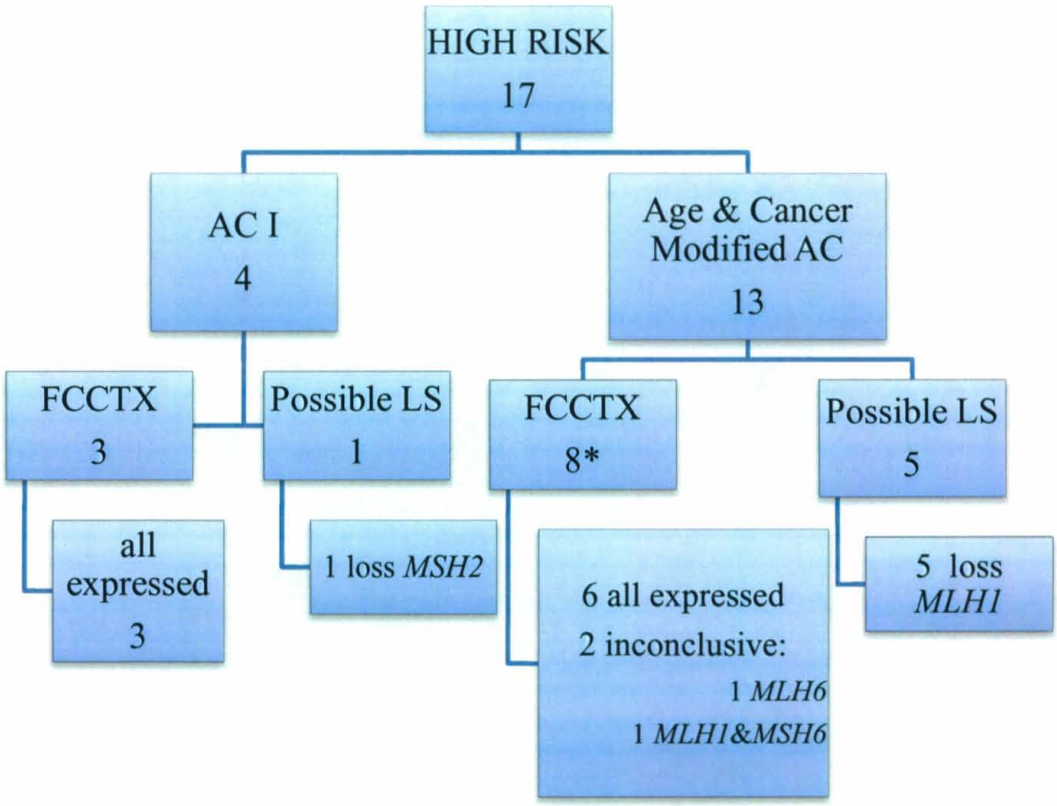
Family history risk classification was completed for one hundred forty-three families (one hundred forty-seven probands) and there is molecular tumour data, both microsatellite and immunohistochemistry analyses, for one hundred thirty-eight families (141 probands). As previously indicated the proband with known FAP was not family history risk classified and therefore was not included in this portion of the results. Tumours with MSI or inconclusive IHC staining are highlighted throughout the tables.

Eight probands required special consideration in this analysis as they represented only four families but each proband had a tumour collected and molecular analysis completed. As all eight tumours were MSS and expressed all three proteins only one tumour representing each family was included for a total of one hundred thirty-nine tumours.

3.4.1 *High Risk Families*

Seventeen (94.4%) family classified high risk families had molecular data to be compared (Figure 4 and Figure 5). Families were classified as possible LS, FCCTX, or age modified FCCTX according to IHC results. As previously noted, without demonstrating a deleterious variant in a MMR gene LS cannot be definitively diagnosed

Figure 4: Summary of Classification of High Risk Families According to Molecular Data Results. Families classified as high risk by family history were further classified according to IHC results. Families were classified as possible Lynch syndrome if there was a loss of protein expression and FCCTX/age-modified FCCTX if all proteins were expressed.



*One tumour had incomplete microsatellite instability data

Figure 5: Molecular Data Correlation According to Specific High Risk Category. Families initially received risk classification according to family history. High risk families fulfilled either ACI criteria or ACMAC. Molecular analysis of tumour protein expression further classified these families as possible LS (loss of protein expression) or FCCTX (no loss of protein expression). Particular protein expression is highlighted according to family classification.

but also cannot be excluded. While IHC is not a surrogate for determining LS it does raise clinical suspicion for a hereditary syndrome. Six (35%) families were diagnosed with possible LS and demonstrated MSI. One family fulfilled the ACI criteria and had loss of expression of *MSH2*. The remaining five families fulfilled ACMAC and were deficient for *MLH1*.

Three (18%) families were classified as FCCTX while seven (41%) were age modified FCCTX. Eight of these families expressed all proteins while two had inconclusive results. Seven families were MSS with one tumour having incomplete microsatellite results.

One tumour from a patient classified as ACMAC could not be collected and was not included in the combined analysis.

3.4.2 *Intermediate Risk Families*

Forty-seven cases (97.9%) of those classified as IR had molecular data available (Figure 6 and Figure 7). Forty-six had complete microsatellite and immunohistochemistry results with one family fulfilling Bethesda criterion 5 missing microsatellite data. Ten families fulfilling Bethesda criterion 1 were all MSS with positive expression of all proteins. Of the six Bethesda criterion 2 families, five were MSS and positive for all proteins, and one was MSI with loss of expression of *MLH1*.

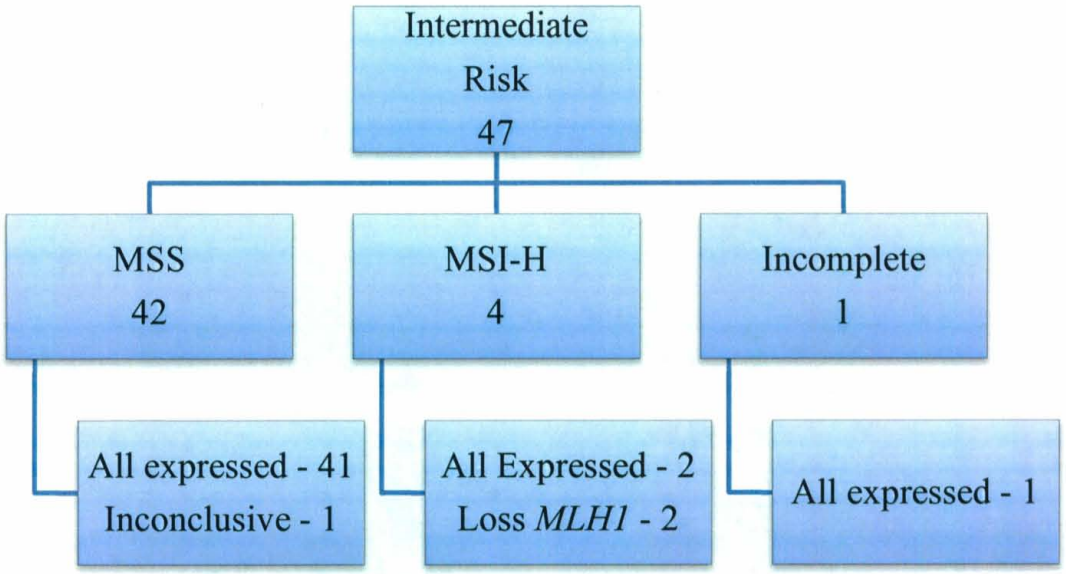
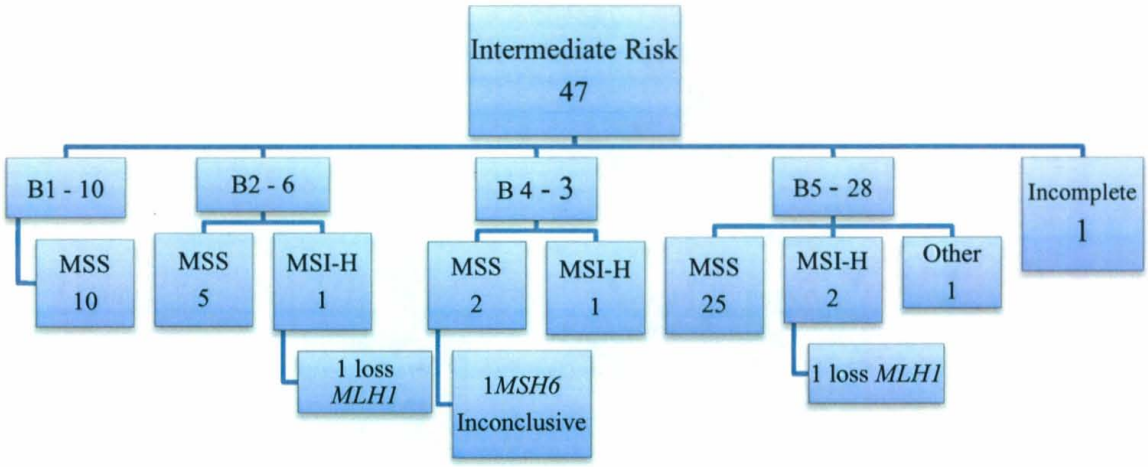


Figure 6: Molecular Data Correlation of MSI and IHC for Family History Classified Intermediate Risk Families



NOTE: Unless otherwise specified, tumours demonstrated positive expression for all proteins

Figure 7: Molecular Data Correlation of MSI and IHC for Family History Classified Intermediate Risk Families According to Individual Revised Bethesda Guidelines. Intermediate risk families fulfilling more than one Bethesda criteria were classified according to the criteria most suggestive of a Mendelian inheritance pattern. Molecular data results were then correlated with each specific guideline.

Two of the three families meeting Bethesda criterion 4 were MSS with one expressing all proteins and one inconclusive for *MSH6*. The third family was MSI and positive for all proteins. Twenty-seven of the twenty-eight Bethesda criterion 5 families had complete data with twenty-five demonstrating MSS and expression of all proteins. Two families were MSI with one expressing all proteins and one with loss of expression of *MLH1*. The one family with incomplete microsatellite data expressed all proteins.

3.4.3 ***Low Risk Families***

Seventy-four low risk families had complete molecular data available with one tumour unable to be collected (Figure 8).

A total of forty-five informative families were MSS of which forty-three were positive for all proteins and two had inconclusive results for *MSH6*. Four families were MSI with three having loss of expression of *MLH1* and one with a loss of *MSH6* and *MSH2*.

Twenty-five non-informative low risk families had molecular data available. Twenty-four were MSS with twenty-three expressing all proteins and one inconclusive for expression of *MLH1*. One family was MSI with loss of expression of *MLH1*.

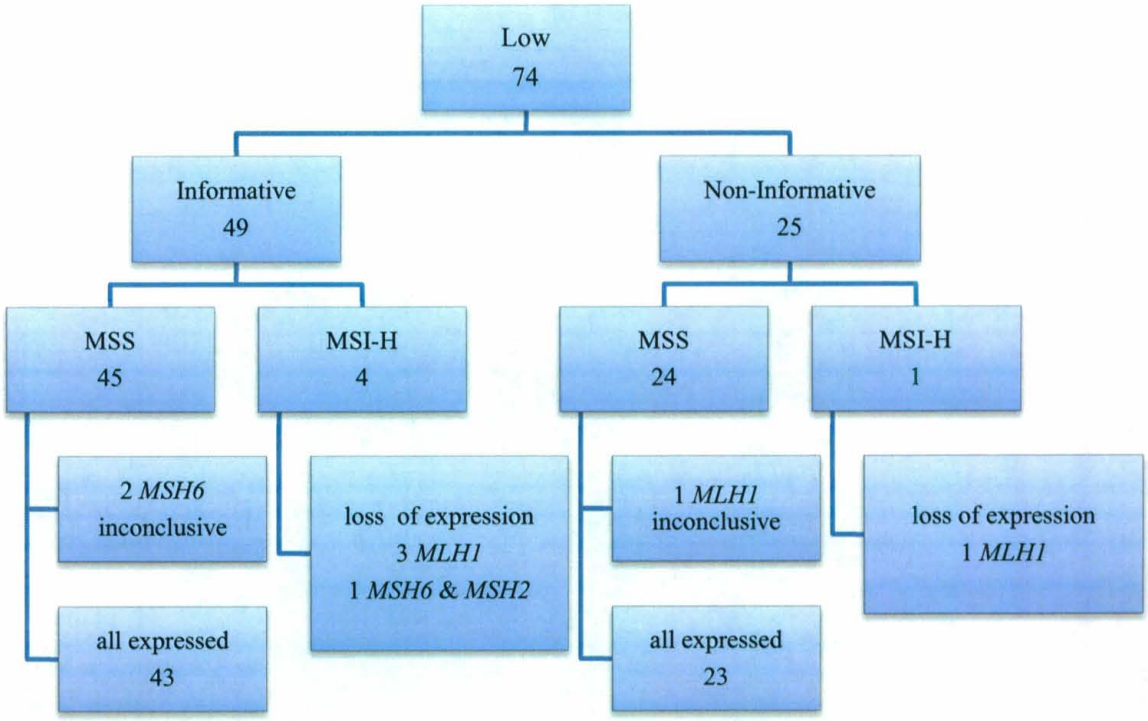


Figure 8: Family History and Molecular Data Correlation of MSI and IHC for Family History Classified Low Risk Families

3.5 *Phenotype Considerations*

First and second degree relatives of families classified as possible LS or FCCTX were compared according to age of onset and frequency of CRC and ACII cancers. Kaplan Meier survival analysis was used to illustrate any differences in age of onset. There were a total of seventy-eight family members representing possible LS and one hundred fifty-five for FCCTX.

3.5.1 *Colorectal Cancer*

Possible LS families had a mean age of CRC onset of fifty-six years of age. The minimum age was twenty-three and the maximum eighty-six. The frequency of occurrence was 24%. FCCTX families had a mean age of CRC onset of sixty-two years of age with a minimum age of thirty-eight and maximum of eighty-nine years. Frequency of CRC was 17% (Table 13). Figure 9 illustrates the later age of onset of CRC in FCCTX families compared to possible LS which was not statistically significant (Mantel-Cox p=0.202, Breslow p=0.101).

3.5.2 *Extra-Colonic Cancers*

As outlined in section 1.2, extra-colonic cancers are a well-defined component of LS. Comparison of family members with cancers as defined by the RBG including colorectal, endometrial, stomach, ovarian, pancreas, ureter, renal pelvis, biliary tract, brain-glioblastoma, small intestine, sebaceous gland adenomas and keratocanthomas was completed. Possible LS families had a minimum age of onset of these extra-colonic cancers of twenty-one years and maximum age of eighty-six years. The average age

Table 13: Comparison of Age of Onset of CRC Occurrence between possible LS and FCCTX Families. Minimum and maximum ages as well as mean age of CRC onset for possible LS and FCCTX family members were compared. Frequency was calculated based on total family members reported by family history.

	Min. Age	Max Age	Mean Age	Frequency
FAMILY				
Possible LS	23	86	56	19/78=24%
FCCTX	38	89	62	27/155=17%

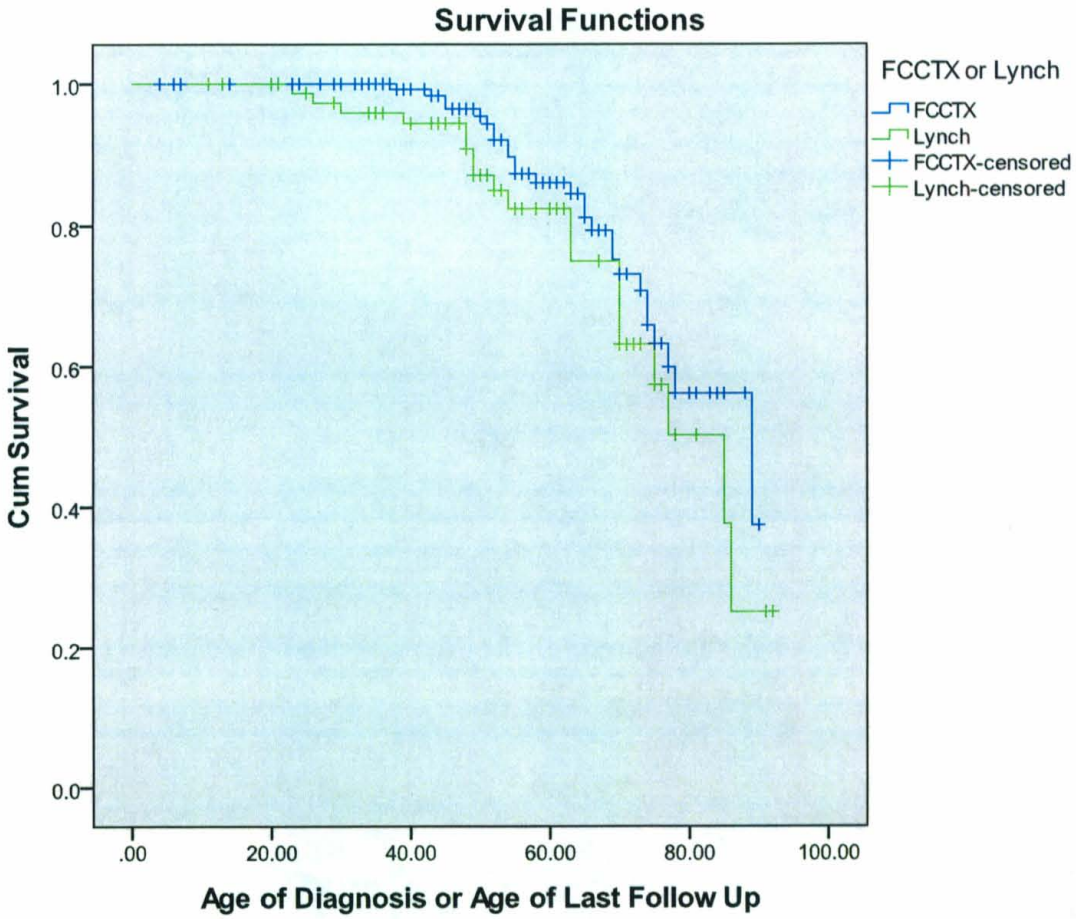


Figure 9: Comparison of Age of Onset of CRC Occurrence between possible LS and FCCTX Families. Kaplan Meier survival curve demonstrates the later age of onset of CRC in FCCTX families compared to possible LS. Age of onset was used for those with a diagnosis of CRC and age of last follow up was used for those without occurrence. (Mantel-Cox p=0.202, Breslow p=0.101).

of onset was fifty-four years. The frequency of these cancers was 35%. This is compared to the FCCTX families that had a minimum age of onset of thirty-eight and maximum age of eighty-nine years old. Twenty three percent of FCCTX family members developed LS associated extra-colonic cancers (Table 14). The Kaplan Meier curve in Figure 10 illustrates the earlier age of onset of these cancers in possible LS families compared to FCCTX (Mantel-Cox p=0.069, Breslow p=0.08).

Table 14: Comparison of Age of Onset of Extra-Colonic Cancer Occurrence between possible LS and FCCTX Families. Minimum and maximum ages as well as mean age of LS associated extra-colonic cancer onset for possible LS and FCCTX family members were compared. Frequency was calculated based on total family members reported by family history.

	Min. Age	Max Age	Mean Age	Frequency
FAMILY				
Possible LS	21	86	54	27/78=35%
FCCTX	38	89	63	36/155=23%

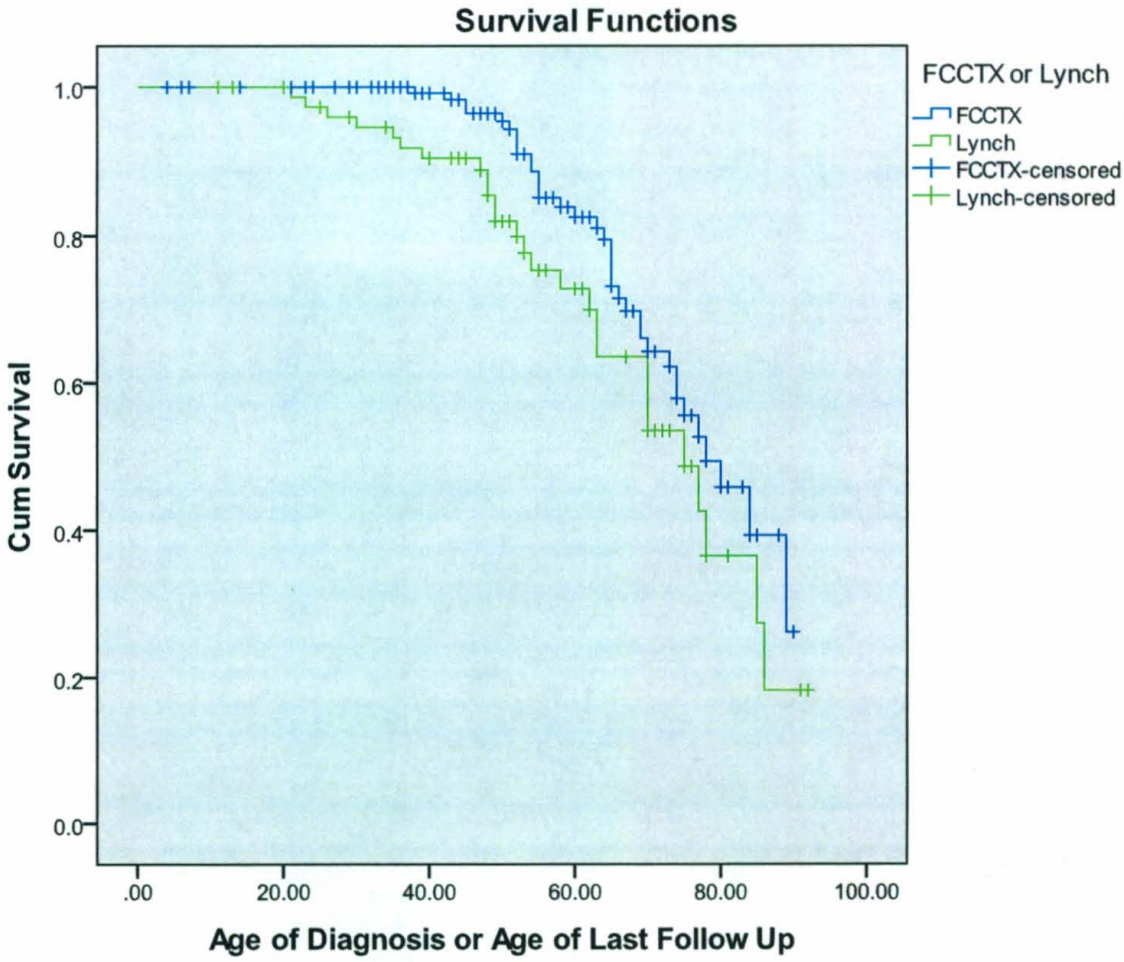


Figure 10: Comparison of Age of Onset of Extra-Colonic Cancer Occurrence between possible LS and FCCTX Families. Kaplan Meier survival curve demonstrates the later age of onset of associated extra-colonic cancers in FCCTX families compared to possible LS. Age of onset was used for those with a diagnosis of an LS associated extra-colonic cancer and age of last follow up was used for those without occurrence.

Chapter 4: *Discussion and Conclusion*

As previously noted, this study was designed to only utilize family history and tumour molecular analyses, not the mutation analysis that was being completed for the entire CRC-IHRT study. As information regarding this analysis has since been published reference to those results are included where applicable.

The incidence of CRC with high or intermediate family history risk is high in Newfoundland and Labrador. 12.5% of probands had a family history consistent with autosomal dominant disease, such that pedigrees show a fifty percent risk that each child of a carrier will inherit the relevant variant and have a high risk of developing colorectal or other associated extra-colonic cancers. Possible LS was diagnosed in a minority of these families, and FCCTX/age modified FCCTX in the majority. Phenotype differences exist between FCCTX and possible LS families in both age of onset, type of cancer, and cancer frequency.

The high rate of familial CRC in NL is readily evident given that greater than forty-five percent of families fulfilled high or intermediate family history risk criteria. Of these, 12.5% were classified as ACI families while there were no ACII families. This is more than double the percentage cited by other population-based studies which have reported incidences between 0.3-5.8% (Aaltonen et al., 1994; Kee & Collins, 1991; Mecklin et al., 1995; Ponz de Leon, 1993; Westlake et al., 1991). While there were no families that met ACII, there were a large proportion of families (77.8%) that displayed multiple members across several generations with CRC and related tumours but failed to meet the age restriction as defined by the Amsterdam criteria. Consequently, another classification known as ACMAC was developed to further delineate this group. Families

who failed to meet the ACII classification based on age alone were captured by adjusting the age of diagnosis from 50 to 60 and including cancers defined by the RBG. The Amsterdam criteria have been criticized for the stringent age restriction and this study provides evidence for the value of this new classification. By increasing the age limit to 60, more high risk families were identified. Consequently, by clinically identifying more high risk families earlier screening and treatment protocols can be implemented for a larger number of at risk individuals that would otherwise have been missed. This high percentage of familial CRC does not seem to be attributable to recruitment bias. No significant differences between participants and non-participants were found when sex, mean age of diagnosis, CRC tumour site, and distribution of residence were compared. Unfortunately, only 40% of cancers affecting family history risk classification were able to be verified. One could therefore argue a falsely elevated Amsterdam classification secondary to inaccurately reported family cancers. However, supporting literature demonstrates a high level of accuracy in self-reported cancers and should not be discounted (Love et al., 1985; Kerber et al., 1998; Sijmons et al., 2000).

Six (35%) high risk families were considered possible LS. Suspected diagnosis was based on the combination of molecular tumour results and high risk family history classification of the proband (family member) included in this study. Specifically, probands classified as high risk and whose tumours demonstrated MSI and MMR protein deficiency. Since the molecular elucidation of LS, the main component for hereditary classification has focussed on the presence of MMR deficiency. More specifically, it has been defined by familial clustering of colon cancer along with evidence of MMR protein

deficiency with *MLH1* and *MSH2* accounting for 80-90% of MMR mutations in LS (Boland, 2005). Of these six possible LS families, one had deficiency in *MSH2* while the other five lost expression of *MLH1*. Interestingly, only one family fulfilled the traditional ACI and lost expression of *MSH2*, while the other five were family classified as ACMAC and lost expression of *MLH1*. As reviewed in section 1.6.1, MMR deficiency is not an exclusive characteristic of LS tumours. 15-20% of sporadic CRC tumours demonstrate MSI and *MLH1* deficiency due to hypermethylation of the *MLH1* promoter region (Poulogiannis et al., 2010). Consequently, once LS is clinically suspected, it is recommended to carry out germline testing of DNA MMR genes for definitive diagnosis (Vasen et al., 2007; Julie et al., 2008; Jang et al., 2010). Woods et al. (2010) tested all tumours for *MLH1* promoter methylation as well as completing DNA sequencing for tumours that were MMR deficient but lacked *MLH1* promoter methylation. For the 1999 cohort discussed in this study, the only deleterious variant reported was for the ACI family which demonstrated *MSH2* deficiency. Consequently, while *MLH1* hypermethylation could explain the *MLH1* deficiency in the remaining five ACMAC families, the possibility of germline variants cannot be excluded. While not common, these ACMAC families could exhibit germline methylation which has been reported in the literature in individuals meeting clinical criteria for LS (Suter et al., 2004; Hitchins et al., 2005). Secondly, of the forty-five reported *MLH1*deficient tumours by Woods et al. (2010), eight had no evidence of methylation, no mutation could be found, and they were not able to determine the occurrence of *MLH1* epimutations to account for the results. While the diagnosis of LS cannot be confirmed, the family characteristics of these five ACMAC *MLH1* deficient families still exhibit significant clustering of CRC and LS

associated extra-colonic cancers. So while they may not have an identifiable hereditary syndrome, they still demonstrate an autosomal dominant inheritance pattern suggesting the possibility of another familial type of CRC.

Interestingly, MMR deficiency was found in an additional seven families that were not family history classified as high risk; two were classified as IR while five were low risk and require further discussion.

Of the two IR families, one fulfilled the revised Bethesda criterion 2, the other criterion 5 but both were deficient in *MLH1*. Using clinical and pathological criteria, the RBG were established to capture a high proportion of LS patients by selecting tumours that should be tested for the presence of MSI. While not diagnostic, MSI is strongly associated with LS as it is usually a phenotypic marker of MMR deficiency, however, for tumours deficient in *MLH1* it does not identify those that are due to hypermethylation (Boland, 2007; Poulogiannis et al., 2010). Woods et al. (2010) reported four mutations for families fulfilling RBG, none of which were *MLH,* which was the protein deficient in my families. As described by Boland & Goel (2010), these families also exhibit the characteristic features of sporadic CRC with MSI including the lack of significant familial clustering and absence of the *MLH1* protein. Consequently, it is highly probable the MMR deficiency demonstrated by these two families is secondary to *MLH1* hypermethylation. Interestingly, while 33% of my 1999 cohort was family classified as IR based on the RBG only 8.5% demonstrated MSI and 4.25% had MMR deficiency. This low specificity of RBG is consistent with other studies lending further support they

may not be as useful as expected for identifying patients whose tumours should undergo molecular testing (Trano et al. 2010; Terdiman et al. 2001; Syngal et al. 2000).

The five low risk families with MMR deficiency included four informative families and one non-informative. In other words, the majority of these low risk families had adequate size pedigrees to demonstrate familial clustering but failed to do which decreases clinical suspicion of a hereditary syndrome. The one family classified as non-informative was deficient in *MLH1* and had a pedigree of four individuals with only two generations represented. The proband was diagnosed at age 52 and a second colorectal cancer was reported in a parent at age 71 years. A pedigree must have six or more family members in two successive generations potentially at 50% risk of inheriting LS surviving until at least age 60 to be considered informative. With only four family members, this family does not meet that definition and has inadequate family history to fulfill any of the clinical criteria despite demonstrating CRC in 50% of the pedigree in successive generations. While this history appears to be suspicious for a hereditary syndrome, confidence in the correct family history low risk classification increases after review of the mutation analysis reported by Woods et al. (2010). Only one germline variant in *MSH6* was reported for low risk families which was not the study number of the proband in this study. As the same criteria for family history low risk classification was used in both studies, this family should also be represented in the low risk families reported by Woods et al. and therefore correct low risk classification for this family is assumed.

The remaining four low risk informative families included three with MMR deficiency in *MLH1* and one deficient in *MSH2* and *MSH6*. The three families with

deficiency in *MLH1* had CRC diagnosed in the included proband between the ages of 62 and 72. Only one family had another colorectal cancer diagnosed at age 69 and none had other ACII or RBG associated cancers diagnosed. Neither of these families were reported to have specific MMR gene mutations and therefore contribute their loss of *MLH1* due to hypermethylation of the promoter region (Woods et al. 2010). As evidenced in these three families, the characteristic features of sporadic CRC with MMR deficiency, as reported in the literature, include the absence of significant familial clustering, absence of *MLH1* and *PSM2* proteins, and the acquired silencing is usually found in older patients (Boland & Goel 2010; Gryfe, 2006).

The final low risk family, deficient in *MSH6* and *MSH2,* had a family pedigree consisting of fourteen family members dispersed over three generations with the proband diagnosed at age fifty. Only three incidences of cancer were reported: pathologically confirmed cancer of the cervix in a sibling diagnosed at age 47, as well as breast and liver cancer in two second degree relatives at unknown ages of diagnosis. Molecularly, this family lost the *MSH2*/*MSH6* MMR complex and had corresponding MSI suggesting LS, but the familial cancers were not LS associated extra-colonic cancers. Interestingly, Woods et al. (2010) reported this family to have a novel gene mutation in *MSH6* and therefore, while unexpectedly, are confirmed to have LS. While discrepancies between molecular analyses and family history have been well documented it is probable the phenotypic expression seen in this family may be unique to this specific deleterious variant. Alternatively, it may be the result of environmental effects or other genetic factors within this family.

For the total population studied LS incidence was 2.1%. Of one hundred forty four families, three had confirmed germline variants; 2 *MSH2* and 1 *MSH6* (Woods et al., 2010). Comparatively, Woods et al. (2010) report a total proportion of 2.7% MMR mutations from a total of 750 patients. This estimate, however, is likely conservative as there were an additional fifteen tumours with MMR deficiency and no identifiable cause. Newfoundland and Labrador demonstrates a similar frequency of possible LS compared to other populations that estimate incidence between 0.9-2.7% (Salovaara et al., 2000; Aaltonen et al., 1998; Cunningham et al., 2001; Ravnik-Glevac et al., 2000; Percesepe et al., 2001; Samowitz et al., 2001; Goodfellow et al., 2001). However, the incidence from our population is likely higher when you include the families with probable novel germline mutations.

There were some inconsistencies between MSI and MMR protein deficiency. MSI is the accepted phenotype of MMR deficiency and although correlation between the two molecular analyses is expected it is not uncommon to find otherwise (Rovella, 2001). Fifteen (10.9%) families demonstrated MSI with only thirteen (86.7%) having a corresponding deficiency in MMR proteins. These two discrepant tumours were from probands who were family history classified as IR; one according to Bethesda criterion-4 and the other criterion-5. The family fulfilling criterion 4 reported ovarian cancer in the proband's sibling at age 35 (permission was declined to review medical records and was unable to be confirmed) and CRC was diagnosed in the proband at age 72. Ovarian cancer is an associated LS extra-colonic cancer and with diagnosis less than age 50 the likelihood of a hereditary component increases. The second family, which fulfilled

Bethesda criterion 5, reported three incidences of cancer: stomach cancer in the proband's parent (pathologically confirmed age 65) and grandparent (age unknown), and lung cancer in the proband's sibling at age 47 (pathologically confirmed). This family has documented ACII cancers in three generations but unfortunately does meet the age criteria for ACII diagnosis. With clear evidence of family history and demonstrated MSI these families are highly suspicious of LS and would be expected to have MMR deficiency. There are several explanations that may account for this. Firstly, this study did not include all known associated MMR genes. Secondly, molecular techniques including discordance between MSI and IHC panels in addition to poor tumour fixation may be responsible (Shia J et al. 2004). Thirdly, research has demonstrated a percentage of LS tumours that do not exhibit any detectable MMR protein expression abnormality despite demonstrating MSI (Poulogiannis et al., 2010; Hampel et al., 2008; Pinol et al., 1986). Finally, a yet to be identified MMR protein that is involved in development of LS may be responsible. Woods et al. (2010) did not report any germline variant for these families and also cited discordance between IHC, MSI, and mutation status.

The majority of family history classified high risk families (65%) were categorized as FCCTX/age modified FCCTX. Specifically, families who fulfilled AC1 and ACMAC but did not demonstrate tumour molecular evidence of LS including MSI or MMR deficiency were categorised as FCCTX and age-modified FCCTX respectively. FCCTX was found in 17.6% and age modified FCCTX in a further 47.1%. Results of mutation analysis for FCCTX families reported by Woods et al. (2010) (they did not include the age modified FCCTX family classification) did not identify any deleterious

variants in the genes extensively studied and there was no evidence of methylator pathway dysfunction for these families. Of note however, three tumours from age-modified FCCTX had inconclusive MMR results. All three were microsatellite stable but had inconclusive staining for *MSH6* which is known to lack the usual phenotypical staining. Specifically, despite having a deleterious variant, *MSH6* is able to preserve some DNA MMR and typical MSI may not be observed and will demonstrate MSS (Kolodner et al., 1999). These ambiguities could be related to poor technique or possibly changes were too subtle to be detected by the usual methods. However, with such a high frequency of these families identified, the more likely explanation is the possibility of novel deleterious variants in this population.

The phenotype of FCCTX (including age-modified FCCTX) families differed from those classified as possible Lynch. FCCTX families had a lower frequency of both CRC and LS associated extra-colonic cancers as well as a later age of cancer onset compared to probable LS family members. With a larger sample size, results from Woods et al. confirmed the age at CRC diagnosis of patients categorised as FCCTX was significantly higher than LS patients, $p<0.011$, however they did not examine the tumour spectrum between FCCTX and LS families. While the higher incidence of cancers for either group may be due to the larger family sizes found in NL and thereby increasing the chance of finding more cancers, these findings are consistent with that found in other studies (Rovella et al., 2000; Park et al., 2007; Perea et al., 2010). Although the results of my cohort were not statistically significant, they may be clinically significant. Insights into phenotypic expression clinically impacts patient management such as necessary

screening protocols for affected family members. Phenotypes are dependent on the underlying genetic background of the mutations in the population studied making direct comparison to other studies difficult. While these observations may suggest a similar phenotypic expression, further information about the particular mutation and genotype-phenotype mechanism is necessary.

There are multiple limitations of this study that must be considered. This study was only one year in duration with a small sample size of 148 participants representing 144 families. Comparative studies have been years in duration with hundreds of participants. A longer study interval, such as the CRC-IHRT which has been recently completed with a total of 750 patients, provides a more reliable data set and reduces the possibility of chance aggregation. Secondly, this was a cross-sectional study with retrospective assessment and verification of reported cancers. If recruitment at diagnosis had been possible it is conceivable that a greater number of relative cancers and tumour samples would have been obtained. Thirdly, this study did not include all known associated MMR proteins. This may explain the discordance between MSI and MMR protein deficiency. *MLH1* and *MSH2* are the most common proteins affected, however, *PMS1* and *PMS2*, albeit to a lesser extent, have also been identified (de la Chapelle, 2004). Additionally, the contribution of *MLH1* hypermethylation was not determined. Up to 10-12% of all CRCs which show MSI develop instability due to this acquired defect while only 3-4% is due to LS (Boland et al. 2008). Consequently, the absolute proportion of LS families is unable to be determined and only a suspected percentage is possible. Further molecular analysis is needed to appropriately classify the tumours with

MLH1 deficiency. The effect of environment and lifestyle on cancer development was also not evaluated. An analysis comparing lifestyle factors between possible Lynch and FCCTX families would eliminate this possible confounding variable. Finally, this study did not include DNA mutation testing which is an integral part to understanding these genetic syndromes and cannot be overlooked. While this information is now available and referenced where possible, it was not designed to be part of this study.

Despite these limitations, this study has some important implications. The high number of clinically diagnosed FCCTX families as well as LS in this population (using mutation results published by Woods et al. (2010) demonstrates the need for appropriate screening in the whole population. There are likely many families yet to be identified and population based screening should be considered. The high frequency of families classified as FCCTX suggests the possibility of other pathways in the development of CRC in the Newfoundland population. There could be more involvement of deleterious variants in oncogenes or tumour suppressor genes as opposed to the traditional mutator pathway. Given the basis of the NL founder population there exists the possibility of novel genes. Finally, the variations in phenotype are important when counselling and screening affected families. As FCCTX families do not appear to follow the typical LS phenotype one cannot follow the prototypical screening guidelines. Screening may need to be tailored according to one's family and the trends illustrated in this study suggesting less extensive and later onset screening. Further in-depth genetic research is needed in this population to further characterize the mutations and the genotype-phenotype mechanism(s) in these high risk families. Consequently, while the scientific relevance of

this study may be questioned due the lack of mutation analysis results, its clinical relevance cannot be overlooked.

In conclusion, the aim of this study was to determine the genetic basis of colorectal cancer in NL, defined by family history and molecular pathology of the CRC in an incident cohort of CRC cases occurring in the population aged 20-74 during one year. Incidence of CRC with high or intermediate family history risk is high in NL. Families with a history consistent with autosomal dominant disease occurred in 12.5%; LS was presumed in a minority of these families, and FCCTX/age modified FCCTX in the majority. FCCTX families are characterized by later onset colorectal and LS associated extra-colonic cancers, as well as lower family occurrence of these cancers. This study was part of a larger, collaborative project that completed a more in-depth analysis of the hereditary cancer syndromes demonstrated here. Further information on genotype, phenotype and underlying genetic process will allow physicians to better identify high risk individuals at an earlier age and provide them and their families with the appropriate screening and treatment modalities. This will help to decrease the morbidity and mortality associated with these genetic syndromes. Finally, the opportunity for novel gene discovery exists as the molecular genetic basis for most of those with familial CRC was not determined.

REFERENCES

http://www.cancer.ca/Canada-wide/About%20cancer/Cancer%20statistics/CW-Provincial%20stats.aspx?sc_lang=en

Aaltonen, L. A., Salovaara, R., Kristo P. (1998). Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med,* 338, 1481-87.

Aaltonen, L. A., Sankila, R., Mecklin, J-P. (1994). A novel approach to estimate the proportion of hereditary nonpolyposis colorectal cancer of total colorectal cancer burden. *Cancer Det Prev,* 18, 57-63.

Aarnio, M., Mecklin, J-P., Aaltonen, L., Nystrom-Lahti, M., Jarvinen, H. J. (1995). Life-time risk of different cancers in hereditary nonpolyposis colorectal cancer syndrome. *Int J Cancer,* 64, 430-433.

Bear, J. C., Nemec, T. F., Kennedy, J. C., Marshall, W. H., Power, A. A., Kolonel, V. M., Burke, G. B. (1987). Persistent genetic isolation in outport Newfoundland. *American Journal of Medical Genetics*, 27, 807-830.

Boland, C. R. (2007). Clinical Uses of Microsatellite Instability Testing in Colorectal Cancer: An Ongoing Challenge. *J Clin Oncol*, 25(7), 754-756.

Boland, C. R. (2005). Evolution of the nomenclature for the hereditary colorectal cancer syndromes. *Fam Cancer*, 4, 211-218.

Boland, C. R., and Goel, A. (2010). Microsatellite Instability in Colon Cancer. *Gastroenterology*, 138, 2073-2087.

Boland, C. R., Koi, M., Chang, D., Carethers, J. (2008). The biochemical basis of microsatellite instability and abnormal immunohistochemistry and clinical behaviour in Lynch syndrome: from bench to bedside. *Familial Cancer*, 7, 41-52.

Boland, C. R., Thibodeau, S. N., Hamilton, S. R. (1998). A National Cancer Institute Workshop on Microsatellite Instability for Cancer Detection and Familial Predisposition: Development of International Criteria for the Determination of Microsatellite Instability in Colorectal Cancer. *Cancer Research*, 58, 5248-5257.

Boland, C. R., and Troncale, F. J. (1984). Familial colonic cancer without antecedent polyposis. *Ann Intern Med*, 100, 700-1.

Cunningham, J. M., Kim, C-Y, Christensen, E. R. (2001). The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas. *Am J Hum Genet*, 69, 780-90.

de la Chapelle, A., and Hampel, H. (2010). Clinical Relevance of Microsatellite Instability in Colorectal Cancer. *J Clin Oncol*, 28, 3380-3387.

Esemuede, I., Forslund, A., Khan, A. H. (2010). Improved Testing for Microsatellite Instability in Colorectal Cancer Using a Simplified 3-Marker Assay. *Ann Surg Oncol.* doi:10.1245/s10434-010-1147-4.

Fuchs, C. S., Giovannucci, E. L., Colditz, G. A., et al. (1994). A prospective study of family history and the risk of colorectal cancer. *N Engl J Med*, 331 (25), 1669-74.

Gazzoli, I., Loda, M., Garber, J. (2002). A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the *MLH1* gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumour. *Cancer Res,* 62, 3925-8.

Giardiello, F. M., Brensinger, J. D., Peterson, G. M. (2001). AGA technical review on hereditary colorectal cancer and genetic testing. *Gastroenterology*, 121, 198-213.

Goodfellow, P. J., Buttin, B. M., Herzog, T. J. (2003). Prevalence of defective DNA mismatch repair and MSH6 mutation in an unselected series of endometrial cancers. *Proc Natl Acad Sci USA*, 100, 5908-13.

Grady, W. M. (2003). Genetic testing for high-risk colon cancer patients. *Gastroenterology*, 124, 1574-1594.

Gryfe, R. (2006). Clinical Implications of our advancing knowledge of colorectal cancer genetics: Inherited syndromes, prognosis, prevention, screening, and therapeutics. *Surg Clin N Am.*, 86, 787-817.

Hampel, H., Frankel, W. L., Martin, E., (2008). Feasibility of screening for Lynch syndrome among colorectal cancer patients. *J Clin Oncol*, 26, 5783-5788.

Hampel, H., Stephens, J. A., Pukkala, E. (2005). Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset. *Gastroenterology,* 129, 415-421.

Hendriks, Y. M., Wagner, A., Morreau, H. (2004). Cancer risk in hereditary nonpolyposis colorectal cancer due to MSH6 mutations: impact on counselling and surveillance. *Gastroenterolo*gy, 127(1), 17-25.

Herman JG, Umar A, Polyak K, Graff JR, Ahuja N, Issa JP, et al., (1998). Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci* , 95(12), 6870-5.

Hitchins, M., Williams, R., Cheong, K., et al. (2005). *MLH1* germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. *Gastroenterology*, 129, 415-421.

Hitchins, M., Wong, J., Suthers, G., et al., (2007). Inheritance of a cancer-associated MLH1 germ-line epimutation. *N Engl J Med*, 356, 697-705.

Jarvinen, H. J., Aarnio, M., Mustonen, H. (2000). Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. *Gastroenterology*, 118, 829-34.

Julie, C., Tresallet, C., Broquet, A., Vallot, C. (2008). Identification in daily practice of patients with Lynch syndrome (hereditary non-polyposis colorectal cancer): Revised Bethesda Guidelines-based approach versus molecular screening. *Am J Gastroenterology*, 103, 2825-2835.

Kane, M. F., Loda, M., Gaida, G. M. (1997). Methylation of the h*MLH1* promoter correlates with lack of expression of h*MLH1* in sporadic colon tumours and mismatch repair-defective human tumour cell lines. *Cancer Res*, 57, 808-811.

Katballe, N., Bentzen, S. M., Christensen, M., Wikman, F. P., Orntoft, T., Laurberg, S. (2001). Role of chance in familial aggregation of colorectal cancer. *British Journal of Cancer*, 84(8), 1084-1086.

Katballe, N., Juul, S., Christensen, M., Orntoft, T. F., Wikman, F., Laurberg, S. (2001). Patient accuracy or reporting on hereditary non-polyposis colorectal cancer-related malignancy in family members. *British Journal of Surgery*, 88, 1228-1233.

Kee, F., and Collins, B. J. (1991). How prevalent is cancer family syndrome? *Gut,* 32, 509-12.

Kerber, R., Neklason, D. W., Samowitz, W. S., Burt, R. W. (2005). Frequency of familial colon cancer and hereditary nonpolyposis colorectal cancer in a large population database. *Fam Cancer*, 4, 239-244.

Kerber, R., Slattery, M. L., Potter, J. D., Caan, B. J., Edwards, S. L. (1998). Risk of colon cancer associated with a family history of cancer or colorectal polyps: the diet, activity, and reproduction in colon cancer study. *Int J Cancer*, 78, 157-160.

Kolodner, R. D., Tytell, J. D., Schmeits, J. L. (1999). Germ-line msh6 mutations in colorectal cancer families. *Cancer Res*, 59(20), 5068-5074.

Lichtenstein, P., Holm, N. V., Verkasalo, O. K., (2000). Environmental and heritable factors in the causation of cancer-analyses of cohorts from Sweden, Denmark, and Finland. *N Engl J Med*, 343, 78-85.

Lindor, N. M., Rabe, K., Peterson, G.M., Haile, R., Casey, G., Baron, J. (2005). Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X. *JAMA*, 293, 1979-1985.

Lindor, N. M., Peterson, G. M., Hadley, D. W. (2006). Recommendations for the care of individuals with an inherited predisposition to Lynch syndrome: a systematic review. *JAMA*, 296, 1507-1517.

Love, R. R., Evans, A. M., Josten, D. M. (1985). The accuracy of patient reports of a family history of cancer. *Journal of Chronic Diseases*, 38, 289-93.

Lynch, H. T., and de la Chapelle, A. (2003). Hereditary colorectal cancer. *N Engl J Med*, 348: 919-932.

Lynch, H. T., Drouhard, T. J., Schuelke, G. S. (1985). Hereditary nonpolyposis colorectal cancer in a Navajo Indian family. *Cancer Genet Cytogenet*, 15, 209-13.

Lynch, H. T., Lanspa, S., Smyrk, T. (1991). Hereditary nonpolyposis colorectal cancer (Lynch syndromes I II): genetics, pathology, natural history, and cancer control, part I. *Cancer Genet Cytogenet*, 53, 143-60.

Mecklin, J. P. (1987). Frequency of hereditary colorectal carcinoma. *Gastroenterology*, 93, 1021-5.

Mecklin, J. P., Jaarvinen, H. J., Hakkiluoto, A. (1995). Frequency of hereditary nonpolyposis colorectal cancer. A prospective multi-center study in Finland. *Dis Colon Rectum*, 38, 588-93.

Mueller, J., Gazzoli, I., Bandipalliam, P. (2009). Comprehensive Molecular Analysis of Mismatch Repair Gene Defects in Suspected Lynch syndrome (Hereditary Nonpolyposis Colorectal Cancer) Cases. *Cancer Res*, 69(17), 7053-7061.

Mueller-Koch,Y., Vogelsang, H., Kopp, R. (2005). Hereditary non-polyposis colorectal cancer: clinical and molecular evidence for a new entity of hereditary colorectal cancer. *Gut* , 54, 1733-1740.

Park, I. J., Kim, H. C., Yoon, Y. S. (2007). Clinicopathological Characteristics of colorectal Cancer with Family History: and Evaluation of Family History as a Predictive Factor for Microsatellite Instability. *J Korean Med Sci,* 22(Suppl), S91-7.

Percesepe, A., Anti, M., Roncucci, L. (1995). The effect of family size on estimates of the frequency of hereditary non-polyposis colorectal cancer. *British Journal of Cancer*, 72, 1320-1323.

Percesepe, A., Borghi, F., Menigatti, M. (2001). Molecular screening for hereditary nonpolyposis colorectal cancer: A prospective, population-based study. *J Clin Oncol*, 19, 3944-50.

Perea, J. (2010). Approach to early-onset colorectal cancer: Clinicopathological, familial, molecular, and immunohistochemical characteristics. *World Journal of Gastroenterology*, 16(29), 3697-3703.

Pinol, V., Castells, A., Andreu, M. (2005). Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer. *JAMA*, 293, 1986-1994.

Ponz de Leon, M., Sassatelli, R., Benatti, P., Roncucci L. (1993). Identification of hereditary nonpolyposis colorectal cancer in the general population. The 6-year experience of a population-based registry. *Cancer* 1993, 71, 3493-501.

Poulogiannis, G., Frayling, I. M., Arends, M. J. (2010). DNA mismatch repair deficiency in sporadic colorectal cancer and Lynch syndrome. *Histopathology*, 56, 167-179.

Ravnik-Glavac, M., Potocnik, U., Glavac, D. (2000). Incidence of germline h*MLH1* and h*MSH2* mutations (HNPCC patients) among newly diagnosed colorectal cancers in a Slovenian population. *J Med Genet*, 37, 533-36.

Rovella, V., Carra, S., Cogliandolo, S. (2001). Familial microsatellite-stable non-polyposis colorectal cancer: Incidence and characteristics in a clinic based population. *Annals of Oncology*, 12, 813-818.

Salovaara, R., Loukola, A., Kristo, P. (2000). Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *J Clin Oncol*, 18, 2193-2200.

Samowitz, W. S., Curtin, K., Lin, H. H. (2001). The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. *Gastroenterology*, 121, 830-38.

Shia, J., Ellis, N. A., Klimstra, D. S. (2004). The utility of immunohistochemical detection of DNA mismatch repair gene proteins. *Virchows Arch*, 445, 431-441.

Sijmons, R. H., Boonstra, A. E., Reefhuis, J., Hordijk-Hos, J. M., de Wale, H. E., Oosterwijk, J. C. (2000). Accuracy of family history of cancer: clinical genetic implications. *Eur J Hum Genet*, 8, 181-6.

Sijmons, R. H., Kiemeney, L. A., Witjes, J. A., Vasen, HFA. (1998). Urinary tract cancer and hereditary nonpolyposis colorectal cancer: risks and screening options. *J Urol* , 160, 466-470.

St. John, D. J., McDermott, F. T., Hopper, J. L., et al. (1993). Cancer risk in relatives of patients with common colorectal cancer. *Ann Intern Med,* 118(10), 785-90.

Suter, C. M., Martin, D. I., Ward, R. L. (2004). Germline epimutation of *MLH1* in individuals with multiple cancers. *Nat Genet*, 36, 497-501.

Syngal, S., Fox, E., Eng, C. (2000). Sensitivity and specificity of clinical criteria for hereditary non-polyposis colorectal cancer associated mutations in *MSH2* and *MLH1*. *J Med Genet,* 37, 641-645.

Trano, G., Sjursen, W., Wasmuth, H. H., Hofsli, E., Vatten, L. J. (2010). Performance of clinical guidelines compared with molecular tumour screening methods in identifying possible Lynch syndrome among colorectal cancer patients: a Norwegian population-based study. *British Journal of Cancer*, 102, 482-488.

Thibodeau, S. N., French, A. J., Cunningham, J. M., Tester, D., Burgart L. J., Roche, P. C., et al. (1998). Microsatellite instability in colorectal cancer: different mutator phenotypes and the principal involvement of hMLH1. *Cancer Res,*58, 1713-8.

Umar, A., Boland, C. R., Terdiman, J. P. (2004). Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst*, 96, 261-8.

Valle, L., Perea, J., Carbonell, P., (2007). Clinicopathologic and pedigree differences in Amsterdam I-positive hereditary nonpolyposis colorectal cancer families according to tumour microsatellite instability status. *J Clin Oncol,* 25, 781-786.

Vasen, H. F. A., Mecklin, J-P., Meera, K. P., Lynch, H. T. (1991). The International Collaborative Group on hereditary non-polyposis colorectal cancer (ICG-HNPCC). *Dis Colon Rectum*, 34, 424-425.

Vasen, H. F. A., Moslein, G., Alonso, A., (2007). Guidelines for the clinical anagement of Lynch syndrome (hereditary non-polyposis cancer). *J Med Genet,* 44, 353-362.

Vasen, H. F. A., Saunders, E. A., Taal, B. G. (1996). The risk of brain tumours in hereditary nonpolyposis colorectal cancer. *Int J Cancer*, 65, 422-425.

Vasen, H. F. A., Watson, P., Mecklin, J-P., Lynch, H. (1999). New Clinical Criteria for Hereditary Nonpolyposis Colorectal Cancer (HNPCC, Lynch syndrome) Proposed by the International Collaborative Group on HNPCC. *Gastroenterology*, 116, 1453-1456.

Vasen, H. F. A., Wijnen, J. Th., Menko, F. H. (1996). Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology*, 110, 1020-1027.

Veigl ML, Kasturi L, Olechnowicz J, Ma AH, Lutterbaugh JD, Periyasamy S, et al. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proc Natl Acad Sci U S A 1998;95(15):8698-702.

Wagner, A., van Kessel, I., Kriege, M. G., (2005). Long term follow-up of HNPCC gene mutation carriers: compliance with screening and satisfaction with counselling and screening procedures. *Fam Cancer*, 4, 295-300.

Watson, P., and Lynch, H. T. (1993). Extracolonic cancer in hereditary nonpolyposis colorectal cancer. *Cancer Res Cancer*, 71, 677-685.

Westlake, P. J., Bryant, H. E., Huchcroft, S. A., Sutherland, L. R. (1991). Frequency of hereditary nonpolyposis colorectal cancer in southern Alberta. *Dig Dis Sci*, 36, 1441-7.

Woods, M. O., Hyde, A. J., Curtis, F. K., Stuckless, S., Green, J. S., Pollett, A. F., . . . Parfrey, P. S. (2005). High frequency of hereditary colorectal cancer in Newfoundland likely involves novel susceptibility genes. *Clin Cancer Res*, 11, 6853-61.

Zhang, L. (2008) Immunohistochemistry versus Microsatellite Instability Testing for Screening Colorectal Patients at Risk for Hereditary Nonpolyposis Colorectal Cancer Syndrome; Part II. The Utility of Microsatellite Instability Testing. *Journal of Molecular Diagnostics*, 10(4), 301-307.