

https://dspace.org/

# Molecular profiling of colorectal cancer within South Africa

**McCabe, Michelle**

**2024**
https://hdl.handle.net/10539/37588

Downloaded from a DSpace repository. https://dspace.org/

# Molecular Profiling of Colorectal Cancer within South Africa

Michelle McCabe

Student number: 0312185Y




A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy

Johannesburg, 2022

Declaration:

I, Michelle McCabe, declare that this thesis is my own work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other institution.



.....

(31 December 2022)

Publications:

The results of this study have been published in the following journal articles:

1. M. McCabe, Y. Perner, R. Magobo, S. Mirza and C. Penny. Descriptive Epidemiological study of South African Colorectal Cancer patients at a Johannesburg Hospital Academic institution. JGH Open 2019 4(3):360-367 (Chapter 2)

2. M. McCabe, P. Perner, R. Magobo, P. Magangane, S. Mirza and C. Penny. Microsatellite Instability assessment in Black South African Colorectal Cancer patients reveal an increased incidence of suspected Lynch syndrome. Scientific Reports 2019 9(1):15019 (Chapter 3)

3. M. McCabe, C. Penny, P. Magangane, S. Mirza and Y. Perner. Left-sided Colorectal Cancer distinct in Indigenous African patients compared to Other Ethnic Groups in South Africa. BMC Cancer 2022 22:1089 (Chapter 4)

Abstract:

There is a global requirement to characterize colorectal cancer (CRC) by molecular subtyping within subpopulations for the assignment of relevant therapies to improve prediction outcomes. Previous CRC studies conducted within South Africa (SA) have mainly been epidemiological, and subtyping limited to immunohistochemistry (IHC) protein expression analysis. The conclusions from these reports provided limited insight and lacked supportive molecular investigations in the development of CRC specifically within our population.

CRC develops through 3 main molecular pathways; i.e. (1) Chromosomal instability (CIN) (75-85%), (2) Microsatellite Instability (MSI) (~15%), and (3) CpG Island Methylator Phenotype (CIMP) (15-20%) pathway. This study descriptively analyses histopathological and molecular information of CRC patients in a 5-year study cohort (2011-2015), by assessing MSI status to ascertain unique features associated with different population groups, particularly within the African population. This study showed a large proportion (37%) of African CRC patients present with early disease onset (<50 years) in comparison to other ethnic group (OEG) patients (15%). Molecular characterization of CRC revealed MSI CRC within African patients occurred at an increased frequency compared to other ethnic groups (15% vs 8%), lacked a BRAFV600E mutation, and the dominant deficient (d) mismatch repair (MMR) profile was dMSH2 and dMSH6, suggesting hereditary Lynch Syndrome (LS) as the dominant pathway of disease development. These results are unique, as international findings demonstrate, 75% of MSI-CRC are sporadic, associated with dMLH1 and BRAFV00E mutations. OEG SA patients however were mostly associated with MLH1 and BRAFV600E mutations, and therefore follow the more well-established sporadic MSI pathway. Further insight gained through universal MSI screening was the ability to differentiate MSI-H versus MSS and MSI-L CRC. MSS/MSI-L CRC categorized by tumour site (left versus right) and ethnicity revealed unique histopathological features associated with left-sided CRC (LCC) in African patients compared to OEG patients. In addition,

a higher proportion of MSI-L LCC was seen in African patients associated with more advanced disease stage and unique molecular and histopathological features. These findings suggest MSI CRC found in African patients is predominantly of a hereditary form, and further variant screening analysis to determine causative germline pathogenic variants are required. MSS CRC particularly within the left colon, was associated with unique histopathological features in the African population group, suggesting an alternative carcinogenic pathway of development when compared to OEG patients. MSI-L CRC also illustrated unique features at this site in African patients, and suggest this to be a completely separate molecular subtype. Deeper molecular characterization by next generation sequencing, including somatic and germline cancer gene screening is required to provide more insight into the different molecular subtypes and pathways in the development of CRC within the SA. This research will therefore direct better clinical management strategies, improving diagnosis, genetic counselling and testing strategies, prognosis, treatment outcomes, survival, and the overall burden associated with the disease within SA.

Acknowledgements:

I would like to acknowledge and thank the following people who formed an integral part and assisted me on my journey in completing this study.

To my supervisor, Dr Yvonne Perner, for your encouragement, patience, inspiration, and knowledge. I thank you for your insight and guidance which has been invaluable in this study, I am forever grateful.

To my supervisor, Dr Clement Penny, I thank you for your willingness to have taken me on as a student, for your continuous support and patience, and helping me believe that which I thought was impossible could be attained. I sincerely appreciate your supervision in this project.

To the staff and students within the Division of Anatomical Pathology, National Health Laboratory services, Johannesburg, thank you for your assistance and support in all aspects of my research, I am especially grateful.

To my colleagues, Dr Rindidzani Magobo, Dr Pumza Magangane, and Dr Sheefa Mirza, who were good friends and always willing to assist and giving guidance in manuscript compilation, I am forever grateful.

Lastly, I would like to thank my family: especially my mom who has always been there supporting me, my brothers and sisters in law, my son Jude for his patience, and the rest of my family and friends for the love and support during this journey.

This research was financially supported by the Wits University (FRC grant), National Health Laboratory Services Research Trust, and the Griffin Trust.

Table of Contents

Declaration..... ii
Publications..... iii
Abstract..... iv
Acknowledgements..... vi
List of Figures..... viii
List of Tables..... ix
List of Abbreviations: ..... xi

Chapter 1:..... 1
Introduction ..... 1
1.1. Global burden of CRC..... 1
1.2. CRC in South Africa..... 3
1.3. CRC pathogenesis..... 6
1.3.1. CRC molecular pathways..... 6
1.3.1.1. Chromosomal instability (CIN) pathway..... 7
1.3.1.1.2. *APC* (WNT/β-catenin) signalling pathway..... 8
1.3.1.1.3. RAS (MAPK signalling pathway)..... 9
1.3.1.1.4. *SMAD* (TGF-β signalling pathway)..... 10
1.3.1.1.5. *TP53* (*TP53* signalling pathway) ..... 11
1.3.1.2. CpG island methylation phenotype (CIMP) pathway..... 11
1.3.1.3. Microsatellite Instability (MSI) pathway..... 14
1.3.1.4 Role of miRNAs in CRC..... 18
1.4. Thesis outline and main objectives..... 19

Chapter 2:..... 21
Descriptive epidemiological study of South African colorectal cancer patients at a Johannesburg Hospital Academic institution

Chapter 3: ..... 30
Microsatellite instability assessment in Black South African colorectal cancer patients reveal an increased incidence of suspected Lynch syndrome

Chapter 4..... 41
Left-Sided Colorectal Cancer Distinct in Indigenous African Patients Compared to Other Ethnic Groups in South Africa.

Chapter 5:
Concluding comments..... 52

References:..... 57

Appendices:..... 69
Appendix A: Human Research Ethics Committee Clearance Certificate..... 69
Appendix B: Plagiarism Declaration..... 71

List of Figures

## Chapter 1:

Figure 1: Estimated incidence rate of CRC diagnosed globally..... 1
Figure 2: Estimated mortality rates of CRC diagnosed globally..... 2
Figure 3: Estimated ASR of incidence and mortality rates of CRC, all ages, worldwide..... 3
Figure 4: Estimated incidence of CRC diagnosed in South Africa..... 4
Figure 5: Estimated mortality rate of CRC in South Africa..... 5
Figure 6: Incidence of SA CRC among males and females..... 6
Figure 7: Young CRC patients showing population group proportions in SA..... 7
Figure 8: Proportions of young CRC patients per population group in SA..... 8
Figure 9: Photographic images of the adenoma to carcinoma tumour progression model..... 11
Figure 10: Summary or CRC molecular pathways/subtypes ..... 16
Figure 11: MicroRNA biogenesis..... 18

## Chapter 3:

Figure 1: Flow diagram illustrating the molecular methodology and frequency results of CRC cases diagnosed between 2011–2012..... 35

Figure 2: Map of South Africa illustrating the distribution of population groups by the different provinces..... 37

## Chapter 4:

Figure 1: Haematoxylin and Eosin (H&E) staining (A-C) and Immunohistochemistry (IHC) (D-F) images of CRC tissue sections..... 45

Figure 2: Electropherograms illustrating instability patterns using the 5-mononucleotide MSI PCR panel ..... 46

## Chapter 5:

Figure 1: Algorithmic approach recommended for the molecular characterization of CRC in SA patients..... 56

List of Tables:

Chapter 2:

Table 1: Descriptive histopathological analysis of colorectal cancer cases diagnosed at Charlotte Maxeke Johannesburg Academic Hospital between 2011 and 2015: Black versus other race groups..... 24

Table 2: Descriptive histopathological analysis of colorectal cancer cases diagnosed at Charlotte Maxeke Johannesburg Academic Hospital between 2011 and 2015: Younger (≤50 years) versus older (>50 years) patients..... 25

Table 3: Descriptive histopathological analysis of colorectal cancer cases diagnosed at Charlotte Maxeke Johannesburg Academic Hospital between 2011 and 2015: Male versus female..... 26

Table 4: Descriptive histopathological analysis of CRC cases diagnosed at Charlotte Maxeke Johannesburg Academic Hospital between 2011–2015: Left versus right-sided colon cancer..... 27

Chapter 3:

Table 1: Descriptive analysis of CRC cases diagnosed at CMJAH between (2011–2015) with Microsatellite Instability status, stratified by CRC subtypes: MSI versus MSS..... 34

Table 2: Descriptive analysis of MSI CRC diagnosed at CMJAH (2011–2015), stratified by ethnic groups: Black versus Other Ethnic Groups (OEGs)..... 36

Table 3. Multiple logistic regression factors at risk for MSI CRC..... 37

Chapter 4:

Table 1: Descriptive analysis of MSS CRC cases diagnosed at CMJAH between (2011-2015): Categorized by site: LCC vs RCC and Ethnicity: Indigenous African (IA) vs Other Ethnic Group (OEG)..... 44

Table 2: Descriptive analysis of MSS and MSI-L CRC cases diagnosed at CMJAH (2011-2015): Categorized by MSI status: MSS vs MSI-L CRC and site: LCC vs RCC..... 47

## List of Abbreviations:

The abbreviations used in this thesis describe the following unless otherwise indicated.

| | |
|---|---|
| 5-FU | 5-fluorouracil |
| AA | African American |
| AGO2 | Argonaute 2 |
| AIDS | Acquired immunodeficiency syndrome |
| AJCC | American Joint Committee on Cancer |
| Ala | Alanine |
| APC | Adenomatous polyposis coli |
| BP | Base pair |
| CD117 | Tyrosine-protein kinase KIT |
| CI | Confidence interval |
| CIMP | CpG Island Methylator Phenotype |
| CIN | Chromosomal Instability |
| CIR | Crohn's-like inflammatory reaction |
| CLR | Crohn's-like inflammatory response |
| CMJAH | Charlotte Maxeke Johannesburg Academic Hospital |
| CMS | Consensus Molecular Subtypes |
| CRC | Colorectal cancer |
| dMMR | Defective mismatch repair |
| DNA | Deoxyribonucleic Acid |
| EMAST | Elevated microsatellite alterations at selected tetranucleotide repeats |
| EPCAM | Epithelial Cell Adhesion Molecule |
| FAM | Fluorescein amidites |
| FDR | False Discovery Rate |
| FFPE | Formalin fixed paraffin embedded |
| GSK3b | Glycogen synthase kinase-3 beta |
| H&E | Haematoxylin and Eosin |
| HDI | Human development indices |
| HEX | Hexachord-fluorescein |
| HG | High grade |
| HIV | Human immunodeficiency virus |
| hMLH1 | Human MutL homologue-1 |

| | |
|---|---|
| hMSH2 | Human MutS homologue-2 |
| hMSH3 | Human MutS homologue-3 |
| hMSH6 | Human MutS homologue-6 |
| HNPCC | Hereditary Non-Polyposis Colorectal Cancer |
| HP | Hyperplastic polyp |
| HREC | Human Research Ethics Committee |
| IA | Indigenous African |
| IHC | Immunohistochemistry |
| KRAS | Kirsten Rat sarcoma |
| LCC | Left colon cancer |
| LG | Low grade |
| LIS | Laboratory information system |
| LLS | Lynch-Like Syndrome |
| LncRNA | Long non-coding RNA |
| LOF | Loss of function |
| LOH | Loss of heterozygosity |
| LS | Lynch Syndrome (LS) |
| miRNA | Micro-RNA |
| MMR | Mismatch repair |
| MSI | Microsatellite Instability |
| MSI-H | High level of microsatellite instability |
| MSI-L | Low level of Microsatellite Instability |
| MSS | Microsatellite stable |
| MYC | Myelocytomatosis |
| NCR | National Cancer Registry |
| NHLS | National Health Laboratory Service |
| OEG | Other ethnic group |
| ORG | Other Race group |
| PCR | Polymerase Chain Reaction |
| PIK3CA | Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha |
| pMMR | Proficient mismatch repair |
| PMS2 | Postmeitotic segregation-2 |
| Polε | DNA polymerase epsilon |
| pri-miRNA | Primary miRNA |

| | |
|---|---|
| RCC | Right colon cancer |
| RISC | RNA induced silencing complex |
| RNA | Ribonucleic Acid |
| SA | South Africa |
| SCF | Stem cell factor |
| SCNA | Somatic copy number alteration |
| SFRP2 | Secreted frizzled-related protein 2 |
| SMAD4 | Mothers against decapentaplegic homolog 4 |
| SRCC | Signet ring cell carcinoma |
| SSA | Sessile serrated adenoma |
| TA | Tubular adenoma |
| TACST1 | Tumour-associated calcium signal transducer 1 |
| TCGA | The Cancer Genome Atlas |
| TGF-β | Transforming growth factor beta |
| Thr | Threonine |
| TIL | Tumour infiltrating lymphocytes |
| TNM | Tumour node metastases |
| TP53 | Tumour protein 53 |
| TVA | Tubulovillous adenoma |
| UTRs | Untranslated regions |
| WNT | Wingless-related integration site |

# Chapter 1: Introduction

## 1.1. Global burden of CRC

The burden of colorectal cancer (CRC) remains high globally, ranking 3rd among the most common cancer diagnosed (lung and breast cancer), with an incidence of 1.9 million cases. CRC is the 2nd leading cause of cancer-related mortality (after lung), 935 000 deaths were reported by the WHO in 2020.[1] (Figures 1 and 2) It is also the 2nd most prevalent cancer globally after lung cancer with 5,2 million people living with CRC within 5 years of diagnosis.[1]

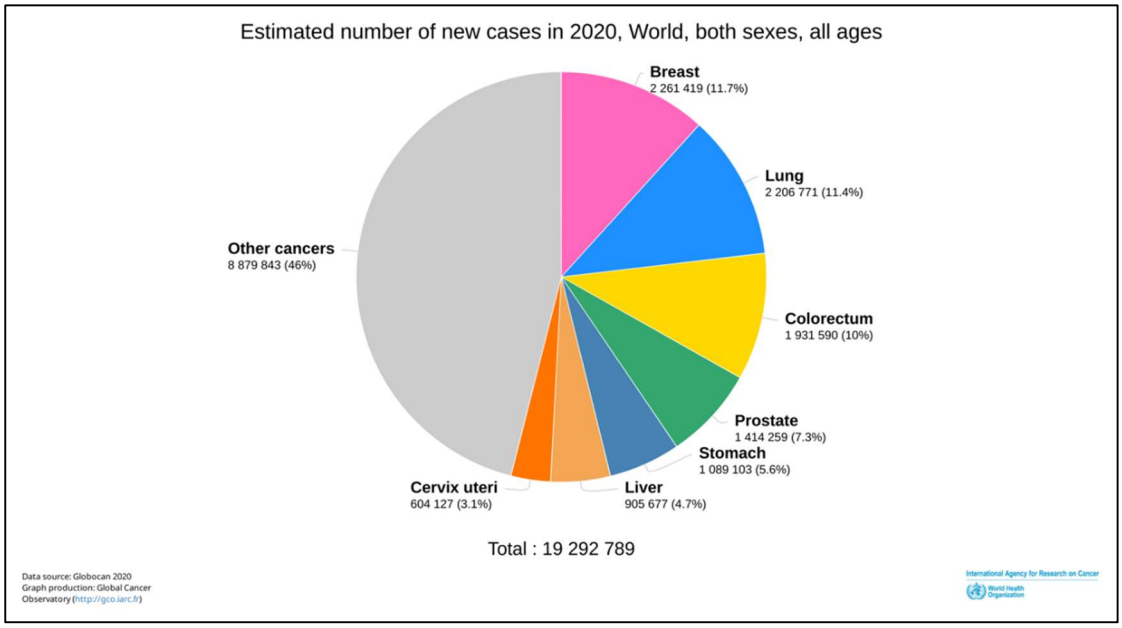



Figure 1: Estimated incidence rate of CRC diagnosed globally. Figure adapted from WHO 2020.[2]

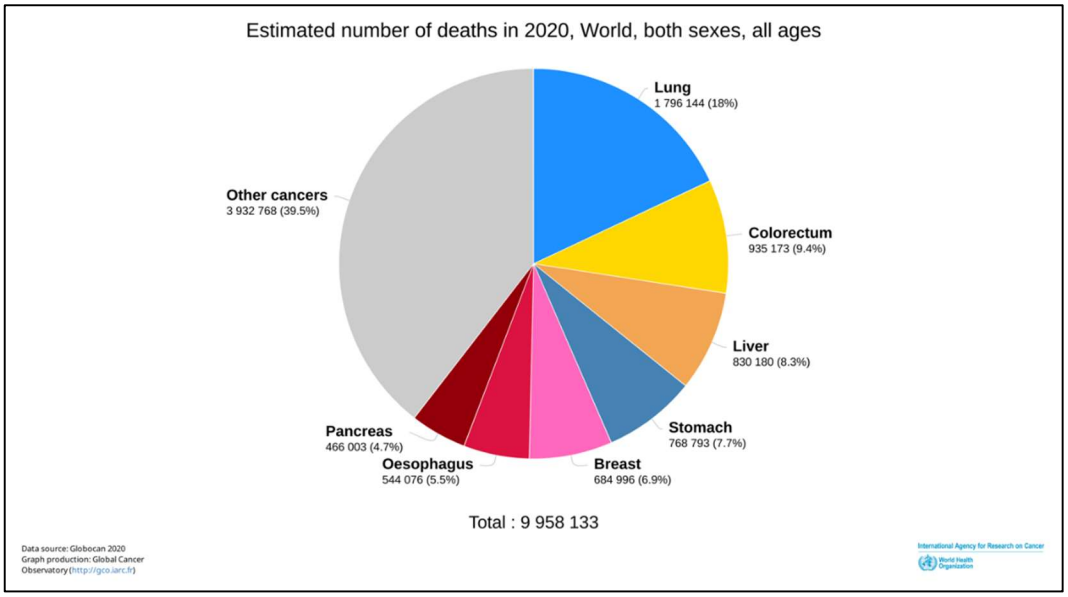



Figure 2: Estimated mortality rate of CRC diagnosed globally. Figure adapted from WHO 2020.[2]

A 10-fold variation in incidence rates exists across the world with very similar patterns in men and women (55% versus 45% respectively) (Figure 3).[1,3–5] The highest rates are reported within the higher human development indices (HDI) or high income countries (i.e. Europe, Australia and New Zealand), whereas the lowest rates were found in the low and medium HDI or low–middle income countries (i.e. South America and Western Asia). [5,6] Even though lower incidence rates are observed in low and medium HDI countries, higher mortality rates of approximately 50-70% are reported reflecting poorer prognosis in these regions.[1,3,5]

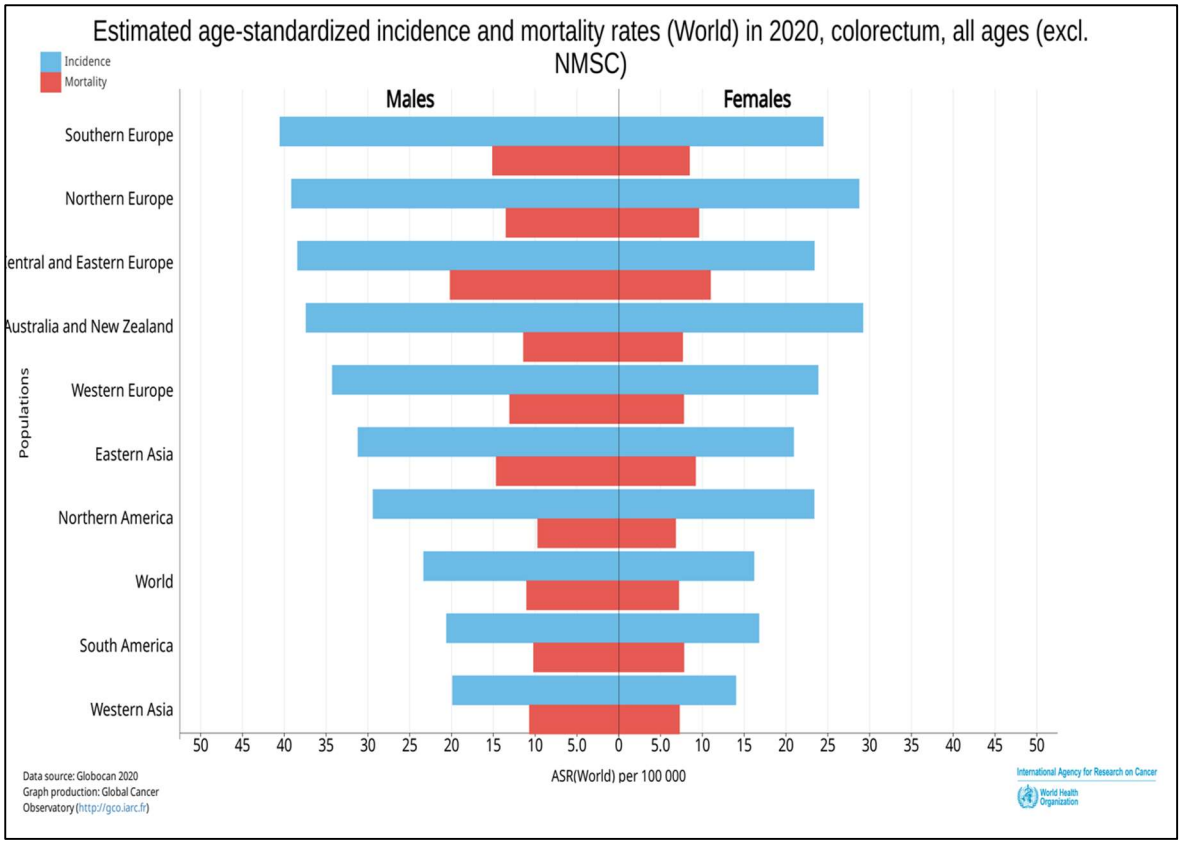



**Figure 3: Estimated ASR of incidence and mortality rates of CRC, all ages, worldwide.** Figure adapted from WHO 2020.[2]

Incidence rates have however shown an increase in historically low risk countries and a decrease in high risk countries.[3,5,7,8] The shift in epidemiology from low to increased incidence is believed to be linked to changes in dietary and lifestyle factors due to urbanization, such as a more sedentary lifestyle, obesity and smoking.[9] In high risk areas such as the United States of America (USA), a decrease in incidence is believed to be the result of more widespread CRC screening for the early detection and removal of precancerous lesions. [7,10,11]

## 1.2. CRC in South Africa:

In Sub-Saharan Africa, South Africa (SA) was found to have the highest incidences of CRC in comparison to the other 19 countries studied in the WHO 2018 report.[9,12] Of the most frequently diagnosed cancers, CRC ranked 5th (Figure 4), 3rd among males (after prostate and lung) and 3rd among females (after breast and cervix).[1] CRC was also identified as the 5th most common cause of death after frequently diagnosed cancers, i.e. lung, cervix, breast and prostate (Figure 5).[1]

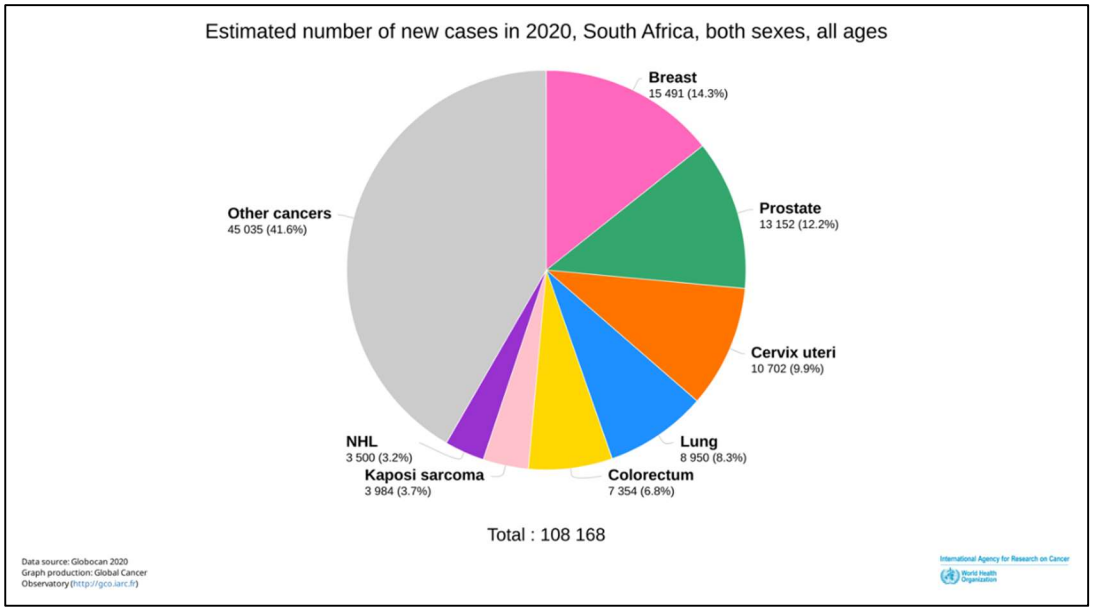



**Figure 4: Estimated incidence of CRC diagnosed in South Africa.** Figure adapted from WHO 2020[2]

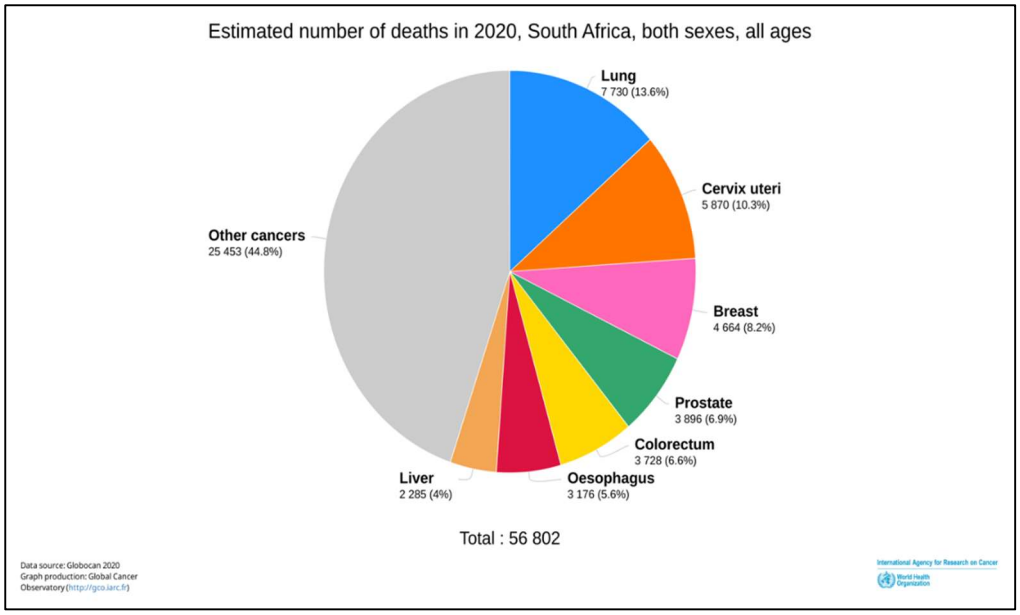



**Figure 5: Estimated mortality rate of CRC in South Africa.** Figure adapted from WHO 2020 [2]

The SA national cancer registry (NCR) largely provides SA's cancer incidence rates, based on annual surveillance reports from histology, cytology and bone marrow aspirate samples analysed from newly diagnosed cancer patients nationally. NCR 2000-2016 reports indicated a steep increase in incidence of CRC over recent years.[13] (Figure 6)

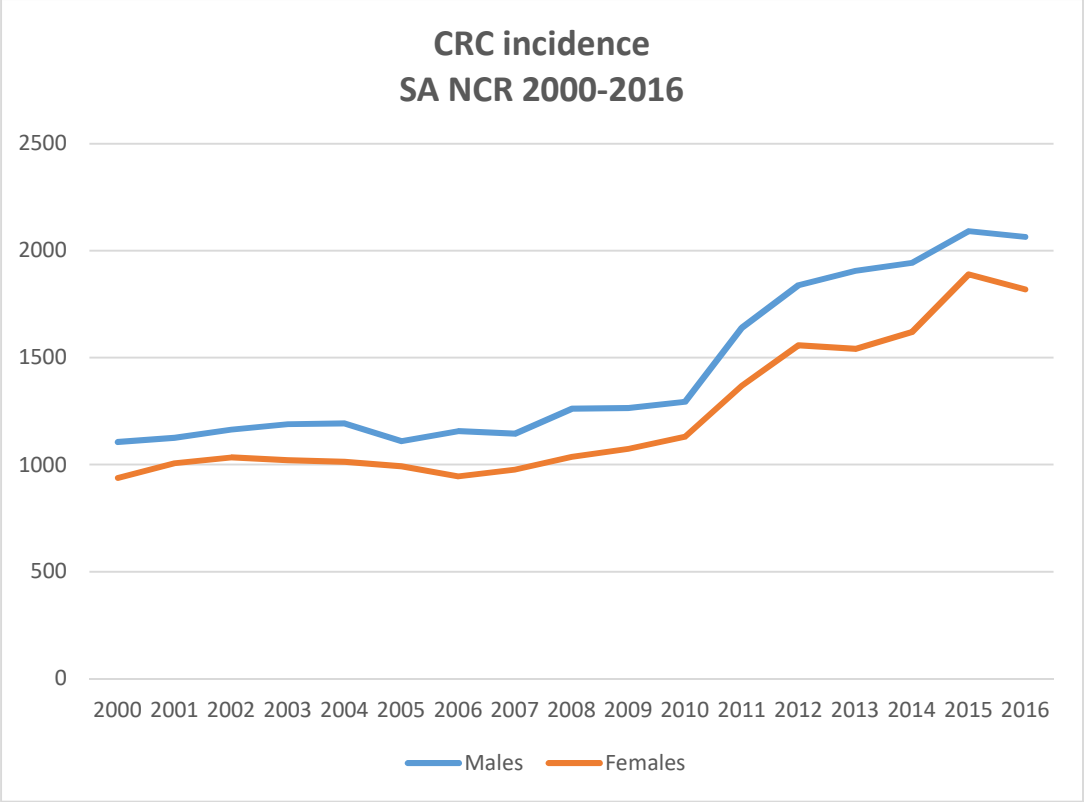



**Figure 6: Incidence of SA CRC among males and females.** Figure adapted from NCR report 2016.[14,15]

There are four main population groups within SA; the dominant group diagnosed with CRC is the white (Caucasian) population group with a frequency of 48-54%. The second is the black (indigenous African) ethnic group with a proportion of 27-31%, followed by the coloured population (mixed ancestry) (13-15%) and the lowest percentages occurring in the Asian (Indian) population group (5-7%). [13–18]

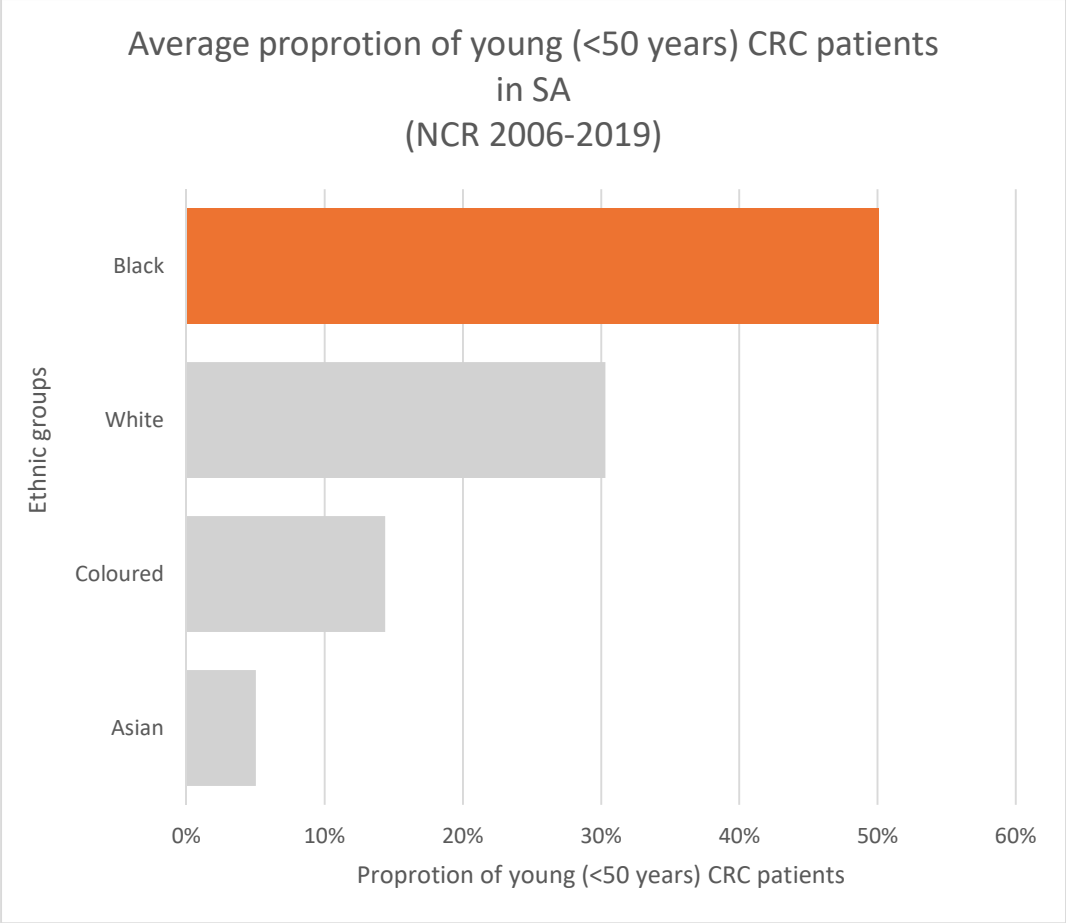



**Figure 7: Young CRC patients showing population group proportions in SA.** Data from NCR 2006-2019.[13,14,24–27,16–23]

Of great concern is the considerably high frequency of young CRC patients (20%) in SA, particularly within the black population group, contributing on average half of all young CRC cases in SA (Figure 7). [13,15,28] This high incidence of young onset CRC particularly in the young black population raises the concern of a possible hereditary contribution. A previous SA CRC study demonstrated a possible familial cancer trend, as approximately 50% of patients had a relative previously diagnosed with cancer, and CRC shown to be the most common.[29]

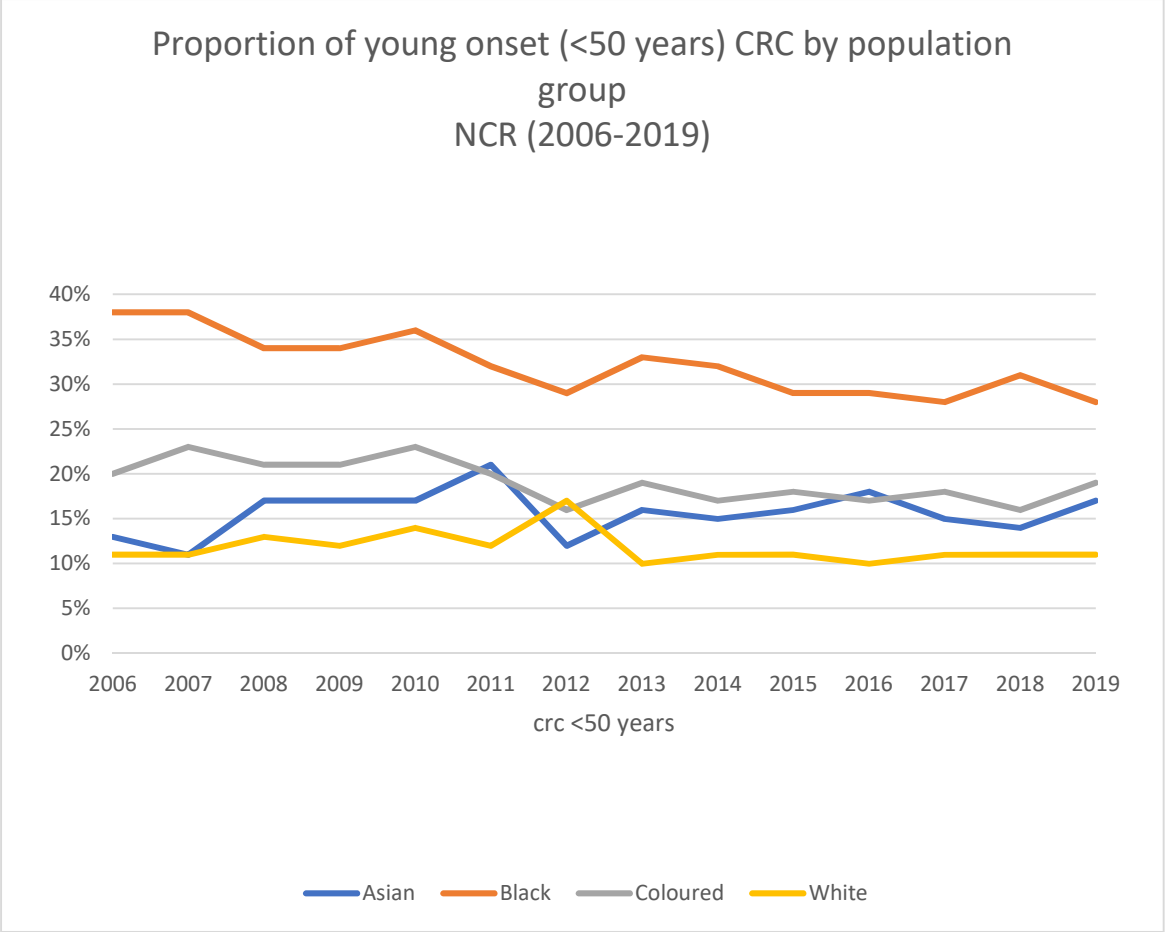



**Figure 8: Proportions of young CRC patients per population group in SA.** Data from NCR 2006-2019.[13,14,23,25–27,15–22]

The trend of young onset (<50 years) CRC within the African population group illustrates a downward trend from 40% to 30% according to NCR 2006-2019 data (Figure 8). The incidence of CRC has increased; this suggests a shifting epidemiology with later age onset (50-65 years) within the black population group. SSA studies consider the shift from rare to rising incidence due to western diet and lifestyle adoption, however no definite aetiological factors have been identified and associated with this trend to date.[9]

## 1.3. CRC Pathogenesis:

CRC is a heterogeneous and multifactorial disease, with the majority (50-60%) of cases due to sporadic changes, such as epigenetic alterations in critical genes, due to unhealthy diet and lifestyle choices.[30] The remainder are a result of familial cancer (20-30%) or genetic (5-10%) predispositions.[30,31] Other studies have also shown bacterial and viral aetiological factors, such as gut microbiota cross-talk in the tumour microenvironment playing a role in the development of CRC. [32–35] These factors lead to a specific sequence of genetic or epigenetic events, transforming healthy epithelial cells to a preneoplastic form and progressing towards more malignant stages as theorized by Fearon and Vogelstein.[36,37] CRC has been shown to arise mainly through 3 different molecular pathways. The majority of CRC (70-80%) develop through the chromosomal instability (CIN) pathway, while the remainder to a much lesser degree follow the microsatellite instability (MSI) (~15%) and CpG island methylator phenotype (CIMP) (15-20%) pathways. [38–42] In addition, regulatory molecules such as non-coding RNAs i.e. microRNA (miRNA), circular (circRNA), long non-coding RNA (lncRNA), packaging RNA (pRNA) and transfer RNA (tRNA), have over recent years illustrated its crucial role in mechanisms of tumorigenesis.[43,44] miRNAs are of particular research interest as they mainly function as inhibitory molecules blocking targeted RNA transcripts from being translated into proteins through transcript degradation.[45]

### 1.3.1. CRC Molecular Pathways:

In 2013, CRC classification by The Cancer Genome Atlas (TCGA) was based on two major molecular subgroups based on genetic instability; (1) MSI accounting for ~16% of cases and the CIN group ~84% of cases. [46] MSI cancers were further characterized into hypermutated (13%) and ultramutated (3%) cancers, with hypermutated cancers occurring as a result of a defective mismatch repair (dMMR) system (~13%) and ultramutated cancer due to pathogenic variants

in the DNA polymerase epsilon (*Polε*) gene. [47] CIN cancer was characterized with aneuploidy and high copy number variation (CNV), however with a microsatellite stable (MSS) genome. Frequently affected genes in this pathway include oncogene Kirsten Rat sarcoma (*KRAS*), and tumour suppressors, Adenomatous polyposis coli (*APC*), Tumour protein 53 (TP53), Mothers against decapentaplegic homolog 4 (*SMAD4*) and Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha (*PIK3CA*).[41,47,48] However in 2015, the Consensus Molecular Subtypes (CMS) Consortium categorized the majority of CRC into four main CMS groups based on meta-analysis CRC expression profiling data. [48] Nearly all hypermutated MSI cancers with prominent immune activity grouped into CMS1 (~14 %), the remainder MSS/CIN cancer subcategorised into CMS2 which follow the canonical *WNT* and *MYC* signalling epithelial pathways constituting the majority (37 %), CMS3 which represent epithelial tumours but with considerable metabolic dysregulation (13%), CMS4 showcase mesenchymal tumours with increased Transforming growth factor beta (TGF-β) signalling (23%), and the residual unclassified group defined by tumours with a heterogeneous phenotype with a combination of features (13%) .[41,48]

#### 1.3.1.1 Chromosomal instability (CIN) pathway

Most CRC develop via the CIN pathway (70-80%), portrayed by gross chromosomal abnormalities (loss of heterozygosity (LOH), insertions, deletions, etc.) A sequence of key molecular events has been well described within this pathway and associated with various stages of tumour development in CRC.[49] Key driver genes include the tumour suppressor genes *APC* and *TP53,* and oncogene *KRAS*, that when altered dysregulate critical molecular signalling pathways leading to cellular transformation [41,50] (Figure 9). The 1st image in Figure 9 illustrates normal mucosa. The first genetic change often detected when cells transition to the next stage of aberrant crypt focus, is due to a loss of function (LOF) mutation in the *APC* gene. This leads

to genomic instability. Additional alterations in the oncogene *KRAS*, allow the cell to proliferate and progress to an early adenoma stage, progressively increasing in size to a late adenoma stage. The final stage of progression includes advanced or high grade (HG) adenoma transitioning to malignant carcinoma and are usually associated with *TP53.* A subset of advanced adenomas may progress due to mutations in *PIK3CA* and chromosome18 LOH.

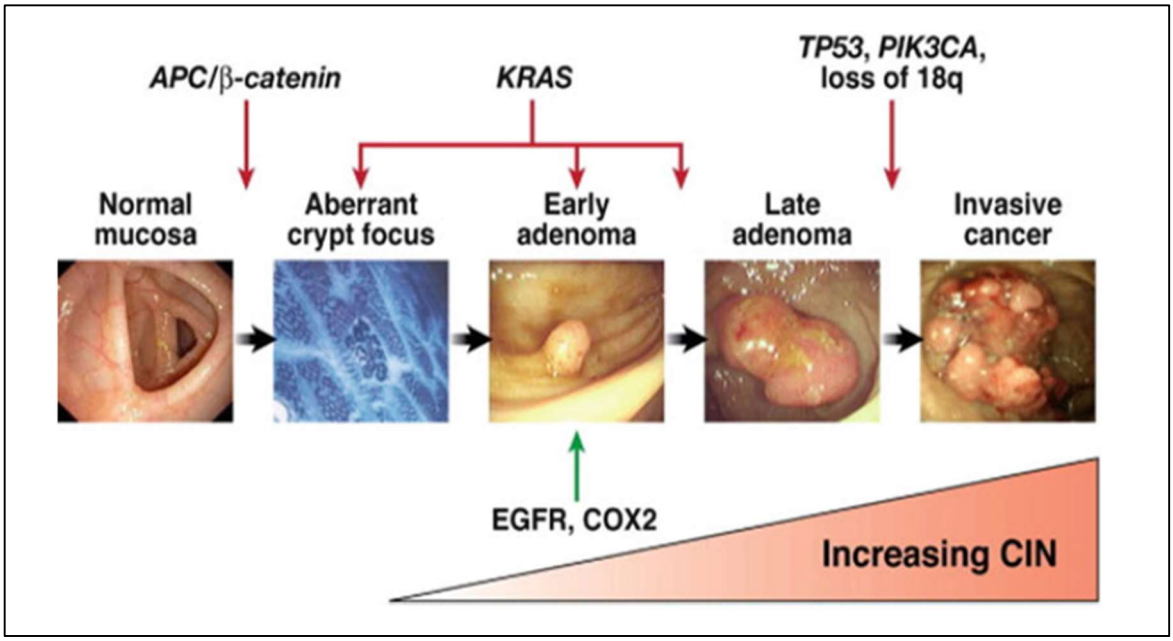



**Figure 9: Photographic images of the adenoma to carcinoma tumour progression model.** Figure adapted from Pino MS, et al (2010)[51]

Patients with CIN type tumours most often have poor prognostic outcome and possess distinct clinical and pathological features in comparison to those with MSI tumours. Determining the molecular and phenotypical features of CRC is essential in ascertaining the appropriate and successful therapeutical regime for successful treatment outcomes. [52]

#### 1.3.1.1.2 APC (WNT/β-catenin) signalling pathway

CIN is mainly linked to the canonical WNT signalling pathway, due to pathogenic variants within the *APC* gene.[41] APC is an important component of the Axin-APC complex, and its main function

is to aid destruction of β-catenin involved in WNT signalling, limiting proliferation in epithelial cells. When *APC* alterations occur and this complex is ineffective, β-catenin is implicated in continuous WNT signalling and the activation and upregulation of important genes such as *MYC* leads to aberrant cell processes and proliferation.[38] In normal colon cells, *APC* is found at increased levels in epithelial cells and decreased levels in base crypt stem cells . The 1st stage of CRC development involves the transition of normal tissue to preneoplastic lesions in aberrant crypt foci cells due to APC LOF, continuous WNT signalling and cell proliferation extending outside crypt baseline cell boundaries (See figure 9).[38] Genomic instability follows this event permitting cells to rapidly progress to an early adenoma stage. Inactive APC are found in sporadic CIN CRC as well as germline mutated *APC* associated Familial Adenomatous Polyposis (FAP).[50]

#### 1.3.1.1.3 RAS (MAPK signalling pathway)

RAS genes are proto-oncogenes that encode for GTPase enzymes involved in the mitogen-activated protein kinase (MAPK) signalling pathway. RAS oncogenes result in the loss of inherent GTPase cleaving activity, leaving GTP in a continuously active form, causing constitutive signalling through the downstream MAPK pathway for continuous cell growth and tumorigenesis. Subsequent events that facilitate tumour progression from a benign to malignant state are initiated by *KRAS* gene alterations. Colorectal tumours harbour up to 50% of *RAS* mutations, with KRAS being the most common mutation (~40%) and *NRAS* least common (2-5%).[53,54] The frequency of *RAS* gene mutations has been found to vary across cancers originating from different tissue types, with the assumption that specific *RAS* isoforms are favoured solely based on the degree of expression. A study by Russo et. al. in 2014 demonstrated increased expression due to altered *KRAS* was mainly associated with colon cancer, whereas altered *NRAS* was seen mostly in rectal cancer. These findings support the theory that *RAS* isoforms having distinct roles in driving cancer from different tissues.[55] *RAS*

mutation status serves as both a prognostic and predictive indicator. Studies have also shown *KRAS* alterations in codon 12 correlated with an enhanced oncogenic phenotype when compared to CRC with mutations in alternative codons such as codon 13. Researchers suggest the latter could therefore only be involved in the initial adenoma-carcinoma transition stage; in contrast to codon 12 variants involved in the advancement of tumour cells to invade and metastasize in the process of carcinogenesis.[56] Codon 12 mutations are therefore associated with poorer prognosis when compared to codon 13 mutations. Screening for *KRAS* and *NRAS* mutations also serve as important prediction markers, as patients with a wildtype genotype for these genes could potentially benefit from anti-EGFR targeted therapy,.[57] Inherited pathogenic variants in *KRAS* and *NRAS* are rarely found and are therefore mostly screened in sporadic CRC.

#### 1.3.1.1.4 *SMAD* (TGF-β signalling pathway)

TGF-β signalling pathway, is a tumour suppressing pathway involving regulation of normal colonic epithelial cell homeostasis in control proliferation, differentiation and apoptosis.[58] In the more advanced tumour stages however, TGF-β signalling has been shown to inversely promote tumour progression.[59] *SMAD4* is a transcription factor in the TGF-β pathway, and it's sequence located on chromosome 18. CRC is often linked to LOH of chromosome 18, which includes the loss of *SMAD4,* and has been found in up to 20% of CRC.[60] Studies have shown disruption of TGF-β signalling leads to more progressive disease within the CIN pathway, and SMAD4 mutation CRC associated with poorer patient survival. The hamartomatous juvenile polyposis syndrome, (JPS), is affected by germline mutations in *SMAD4* or BMPR1A genes and these occur at a frequency of <1% of CRC. [61]

#### 1.3.1.1.5 *TP53 (TP53* signalling pathway)

*TP53* gene alterations are associated with adenoma progression to an invasive malignancy stage.[51] *TP53* is a tumour suppressor gene that is usually only activated in response to DNA damage, stress and oncogene activation.[62] *TP53* has many direct and indirect gene targets (~500 genes) involved in the process of apoptosis and cell cycle arrest.[62] A recent study conducted by Schell et al. 2016, showed pathogenic variants in *APC* , *KRAS* and *TP53* had significantly poorer outcomes when compared to only APC or KRAS mutations, and this could be an important prognostic marker in a routine clinical setting, in predicting outcomes and appropriate therapy.[63] *TP53* germline mutations are associated with Li-Fraumeni syndrome cancers, however CRC are rarer cancers associated with this syndrome.[64]

### 1.3.1.2 CpG island methylation phenotype (CIMP) pathway

The 3rd mechanism (CIMP) is an epigenomic pathway, as a consequence of global methylation of CpG dinucleotide regions in promoter sequences of critical genes throughout the genome.[42] This results in silencing of crucial tumour suppressor genes involved in important cell processes such as the regulation of the cell cycle, DNA repair and apoptosis..[39] CIMP features have been indicated in a fraction of CRC (20-30%) and the associated clinicopathological features found to be similar and overlapping with the MSI phenotype, which include late age onset, female sex, proximal tumours, *BRAF* and *PIK3CA* mutations, and Fusobacterium nucleatum.[65,66] CIMP tumours were also found to be associated with wild-type *APC*, *KRAS* and *TP53* genes, indicating CIMP to be independent of the canonical *WNT* signalling pathway.[65]

### 1.3.1.3 Microsatellite Instability (MSI) pathway:

Fifteen to sixteen percent of CRC is attributed to the MSI pathway caused by disruption of the DNA Mismatch Repair (MMR) system. The functional MMR system is composed of 4 subunits, i.e. human MutS homologue-2 (h*MSH2*), -6 (h*MSH6)*, human MutL homologue-1 (h*MLH1*) and Post-meiotic segregation-2 (*PMS2*). Approximately 12-13% of MSI CRC are the result of somatic alterations (sporadic form), due to methylation silencing of the promoter sequence of *MLH1*, often a feature in older patients.[67] Sporadic MSI CRC also harbour the *BRAF V600E* oncogenic variant as well as clinical and histopathological features similar to CRC following the CIMP pathway. The remaining subset (~3%) are triggered by germline pathogenic variants in either of the 4 MMR genes or the Epithelial Cell Adhesion Molecule (*EpCAM)* gene causing Lynch syndrome (LS), previously referred to as Hereditary Non-Polyposis Colorectal Cancer (HNPCC).[68] LS usually occurs in younger patients.[69] Heritable epimutations although rare, have been described in defective *MLH1* and *MSH2* genes in LS families.[70–72] Methylation of *MSH2* has been reported to occur as a second hit in defective *MSH2* LS patients, and germline mutations in the *EPCAM* gene (formerly called Tumour-associated calcium signal transducer 1 [*TACSTD1]*), located adjacent to *MSH2*, has also been associated with MSH2 heritable somatic epimutations.[73,74]

MSI screening also serves as a useful genetic marker when predicting CRC patient prognosis and survival and assisting clinicians in determining the best patient treatment strategies. MSI is a well-established marker of favourable prognosis and has been promoted to be utilized as a routine test to assist in establishing more accurate prognosis and pathologic staging. [5] MSI is a negative predictor in therapy decision making as MSI tumours have been shown to be resistant to 5-fluorouracil (5-FU) chemotherapy.[2] *POLE* germline mutations have been related to increased risk for inherited CRC and adenomatous polyposis syndromes, particularly when MMR variants have not been identified.[9]

A small subset of MSI tumours have been found with a lack of pathogenic variants in MMR genes, however numerous microRNAs (miRNAs) that potentially inhibit MMR genes transcript from being translated have been found overexpressed in MSI-H CRC, e.g. miRNA-155 and miRNA21.[11,12] The miRNA-155 has been found to downregulate MMR transcripts *MLH1*, *MSH2* and *MSH6*, and miRNA-21 targets the *MSH2* and *MSH6* transcripts.

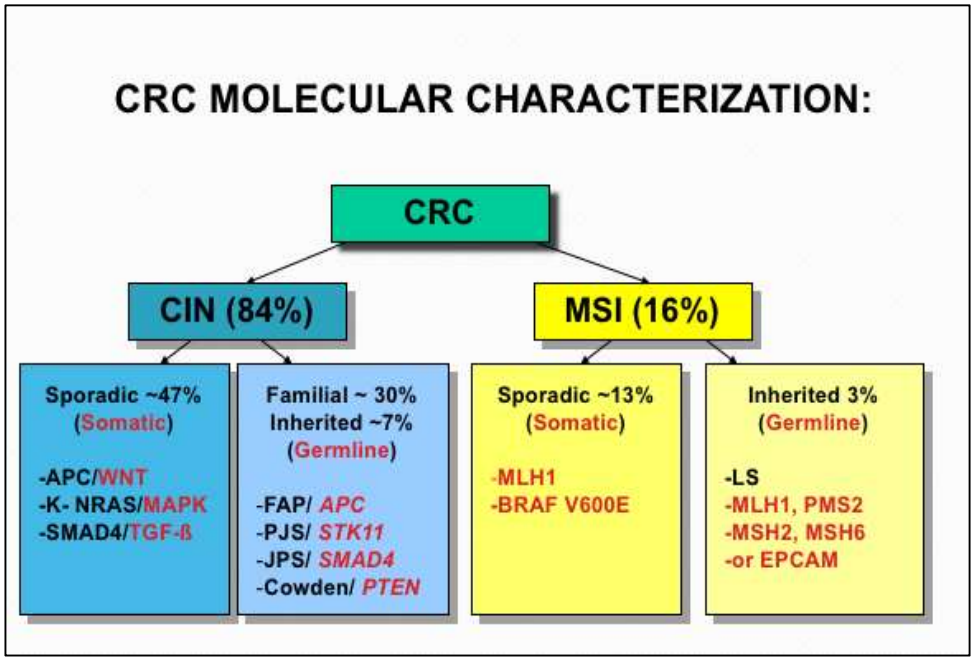



**Figure 10: Summary or CRC molecular pathways/subtypes.** Majority of CRC develop through the CIN pathway. Approximately 30% of CRC is due to a familial component and ~10% inherited. LS contributes to the majority of hereditary CRC cases (3%). FAP-Familial adenomatous polyposis (FAP), Peutz-Jeghers syndrome (PJS) and Juvenile polyposis syndrome (JPS) are the main CIN pathway hereditary syndromes.

### 1.3.1.4. Role of miRNAs in CRC

Approximately 3-4% of the human genome comprises of MicroRNA (miRNA) sequences. MiRNAs are non-coding RNA molecules ranging between 18 and 23 nucleotides in length. They are located within intergenic, intronic or untranslated regions of the genome and are regulated

by protein coding genes.[75] These molecules are critical key players in developmental and biological processes, including cellular differentiation, proliferation, signal transduction, apoptosis, etc.[76] These molecules bind to 3' untranslated regions (UTRs) of target mRNA in the majority of cases, leading to downregulation of gene expression.[77] Literature has also shown miRNAs bind to the 5' UTRs suppressing gene expression, however reports of miRNAs binding to gene promoter regions resulting in transcript upregulation have also been shown.[77–80] A single miRNA has been shown to regulate multiple transcripts, and over 2000 miRNAs have been discovered to date, regulating approximately a third of protein coding genes.[81,82]

miRNAs are also found in clusters, where miRNAs (~2-3) located within close proximity of each other are transcribed together and each have their own specific target.[83,84] miRNAs have been reported to be secreted into extracellular fluid i.e. plasma, serum, saliva, urine, tears etc. and therefore serve as potential cell free biomarkers in disease diagnosis and prognosis.[85–90]

MiRNA synthesis occur via canonical and non-canonical pathways in the nucleus (See figure 11).[75] The canonical pathway occurs through transcription of the host gene, producing a primary miRNA (pri-MiRNA) transcript. These molecules are spliced, capped, and polyadenylated, similar to mRNAs. Cleavage occurs by the Drosha enzyme producing a precursor miRNA, which subsequently moves into the cytoplasm. An endonuclease Dicer, then cleaves the top or loop region of this precursor molecule generating the mature miRNA strand. The mature miRNA strand is then added to the RNA induced silencing complex (RISC), which silence and prevents mRNA targets being translated into protein products. The alternative strand released after Dicer cleavage usually is degraded, however in some cases both strands of the miRNAs may be loaded into RISC, and in these instances strands at the 5′ end of the loop are denoted as "5p" and the 3′ end as "3p". [75] There are alternating non-canonical biogenesis pathways which includes bypassing the Drosha cleavage step, where instead a splicing process generates the precursor. The second is where a miRNA precursor is directly bound and cleaved

by Argonaute 2 (AGO2) instead of Dicer, also resulting in a mature miRNA. The subsequent steps remain the same for each alternating pathway. [75,82]

KRAS has been shown to be inactivated by miR-143, preventing CRC progression, while another study showed miR224 to increase WNT/ß-catenin signalling by deregulation of Glycogen synthase kinase-3 beta (GSK3b) and Secreted frizzled-related protein 2 (*SFRP2*) to translocate ß-catenin in CRC.[91,92] Many studies in CRC illustrating potential roles of miRNA molecules in the downregulation and upregulation of oncogenes, therefore serving as promising tools for future treatment interventions.[84,93,94]

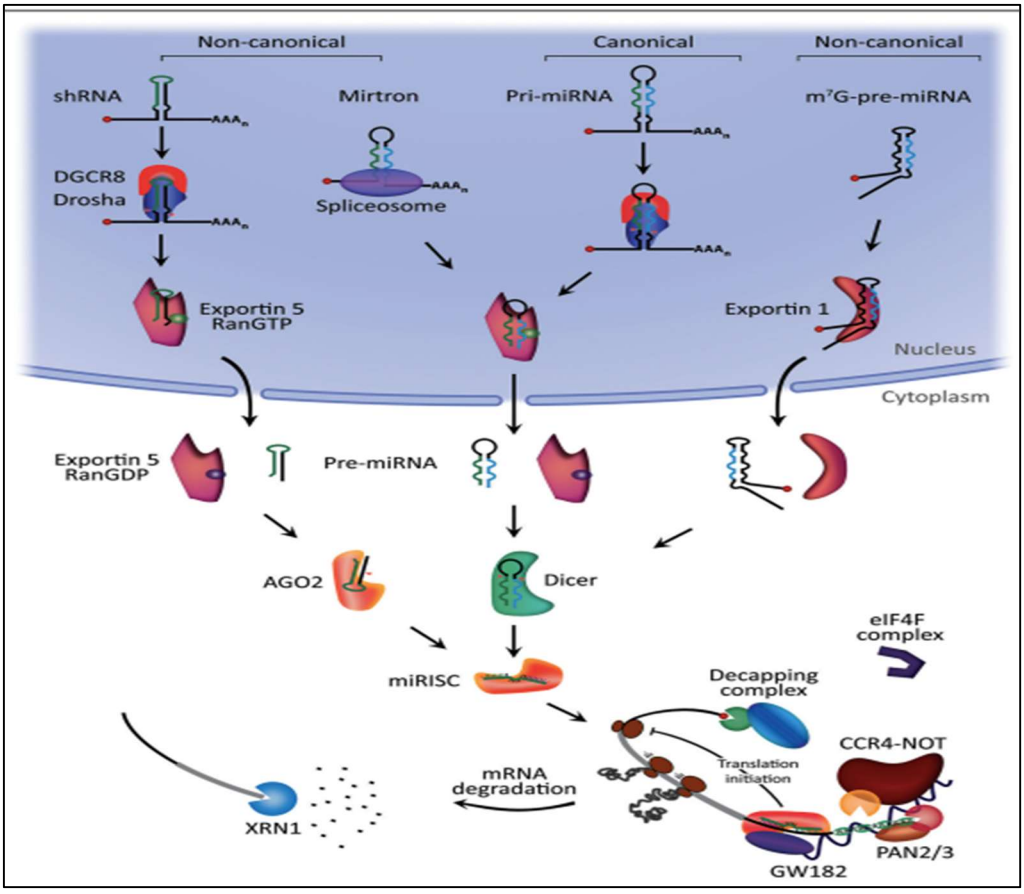



**Figure 11: MicroRNA biogenesis.** The canonical pathway involves the initial step of pri-miRNA transcript generation, followed by the Drosher enzyme complex cleaving it into a precursor-miRNA(pri-miRNA) molecule. It is then exported to the cytoplasm (through Exportin5/RanGTP) and cleaved by Dicer to produce the mature miRNA molecule. Subsequently it is inserted into AGO2 to form miRISC. Figure adapted from O Brien J, et al (2018)[75]

### 1.3.1 Thesis outline and main objectives

Presented in this thesis are the main research findings generated by this study, which includes a deeper understanding of CRC subtypes in SA patients characterized by histopathological and microsatellite instability assessment. These are shown in the form of three original research published journal articles. Chapter 2 includes the published descriptive epidemiological study of CRC patients diagnosed at the Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) over a 5-year period (2011-2015). The results from this study provided insight into the number of patients diagnosed, dominant sex, age, site, ethnicity and histopathological associations seen in the SA CRC population. The findings confirmed the alarming rate of approximately 40% of CRC patients within the African population group are < 50 years of age, compared to ~10% of patients of other ethnic groups. While results from this study identified clinicopathological features associated with site and young age, unique histological features associated specifically with African CRC patients was not seen. This led to our second objective to further characterize by molecular subtyping through MSI screening of all patients diagnosed within a 2-year study period (2011-2012). This was included in our second publication (Chapter 3), where our research showed 12% of CRC patients had a MSI subtype. With additional categorization by ethnicity, interestingly African patients revealed an increased frequency of MSI CRC (15%) when compared to other ethnic groups (8%). In addition, unique associated features included increased MSH2/MSH6 protein expression loss, and the absence of BRAF mutations, suggesting the hereditary form of disease (LS) could be the main cause of MSI CRC in the African population.

The final published journal article characterized the non-MSI CRC group (MSS/MSI-low) to determine whether distinct associations could be identified subsequent to omitting MSI-H patient samples. These findings are presented as an original research article in Chapter 4. MSS and MSI-low samples were descriptively analysed by categorization of ethnicity (African versus

other ethnic groups (OEG) and site (right-sided colon cancer (RCC) versus left-sided colon cancer (LCC). Additional investigations of African CRC categorized by site and MSI subtype revealed associations with gender, site, histological and molecular subtype. These results support the concept that MSI-L CRC possibly is a standalone molecular subtype with a worse prognosis, compared to counterparts MSS and MSI-H CRC.

In conclusion, Chapter 5 summarizes the major findings of our study and suggest prospective research in the much needed understanding of the molecular mechanisms involved in CRC development in SA.

# Chapter 2:

Descriptive epidemiological study of South African colorectal cancer patients at a Johannesburg Hospital Academic institution

JGH Open Open Access An open access journal of gastroenterology and hepatology

JGHF Journal of Gastroenterology and Hepatology Foundation

doi:10.1002/jgh3.12248

ORIGINAL ARTICLE

# Descriptive epidemiological study of South African colorectal cancer patients at a Johannesburg Hospital Academic institution

Michelle McCabe,* Yvonne Perner,* Rindidzani Magobo,* Sheefa Mirza† and Clement Penny†

Departments of *Anatomical Pathology, School of Pathology, Faculty of Health Sciences and †Internal Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

**Key words**
colorectal cancer, descriptive epidemiology, race, South Africa.

Accepted for publication 4 August 2019.

**Correspondence**
Michelle McCabe, Department of Anatomical Pathology, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Parktown, Johannesburg 2193, South Africa.
Email: michelle.mccabe@nhls.ac.za

**Declaration of conflict of interest**: None.

**Funding support:** Griffin TrustGTCP016

## Abstract

**Background and Aim:** Epidemiological studies of colorectal cancer (CRC) in South Africa (SA) have been poorly characterized. Black and white SA population groups have demonstrated distinct CRC clinical presentations, suggesting that black SA patients follow a different carcinogenic pathway than their white counterparts. Thus, the aim of this study was to identify unique demographic and histopathological features associated with black SA patients to facilitate earlier diagnosis and to improve disease management.

**Methods:** This preliminary descriptive epidemiological study included 665 retrospective CRC cases diagnosed between the period 2011 and 2015 at the Charlotte Maxeke Johannesburg Academic Hospital. Demographic and histopathological features in black *versus* other race groups (ORG) were compared, and Student's *t*-test, Chi-square, and Fischer's exact tests were used for statistical analysis.

**Results:** Statistical analysis demonstrated that patients with left-sided tumors of invasive adenocarcinoma were predominantly black and male. These patients were considerably younger when compared to ORG (median 56 *vs* 62 years, respectively), $P < 0.0001$. However, no significant propensity for other histological features was illustrated. Polyps were mostly tubular adenomas (51%) and tubulovillous adenomas (TVAs) (44%). TVAs were mostly high-grade lesions ($P < 0.0001$) and associated with left-sided CRC ($P = 0.0325$).

**Conclusion:** These findings verify that black SA CRC patients have an earlier disease onset in comparison to ORG; however, no increased tendency for tumor site, precursor lesion, stage of disease, or gender was evident. Thus, a deeper molecular characterization of CRC is required to understand the underlying causes associated with earlier disease onset in black SA CRC patients.

## Introduction

Colorectal cancer (CRC) is the third most common cancer worldwide and the fourth most common cause of cancer-related mortalities.[1] Within South Africa (SA), it is the fourth most commonly diagnosed cancer, being the second most common among males and the fourth most common among females and the sixth leading cause of cancer mortalities in SA.[1,2] Male and female ratios have remained constant (53–55% *vs* 45–47%, respectively).[2–6] Male patients have a 1.27 estimated lifetime risk of developing CRC, and females have a lower risk of 0.75.[6] CRC incidence rates within SA have progressively increased over the years, with the annual crude incidence in SA CRC male and female patients in 2014 reported as 7.34 and 5.86 per 100 000, respectively.[6] CRC occurs mostly within the white (Caucasian) population group in SA patients (52–54%), followed by black (African) (26–28%), colored (mixed ancestry) (14–15%), and Asian (Indian) patients (4–7%).[2–6] The pathogenesis of CRC is complex and diverse and is influenced by multiple factors, including diet, lifestyle, and genetic predispositions. It is reported that the white patient population group in SA appears to follow the classic Western trend, which presents at a later age, having an association with diet and lifestyle factors, although the molecular pathology has not been extensively investigated.[7,8] Comparatively, the black SA population has a higher frequency of young (<50 years) CRC patients, possibly because of diet and lifestyle changes either due to urbanization or a familial contribution.[7–9] A study by Prodehl *et al.* showed that 49.0% of SA CRC patients reported a family history of cancer, with CRC the most frequently diagnosed, where maternal (35%) and fraternal (33%) members were the most affected, followed by paternal members (25%).[10]

High prevalence rates for young CRC patients, <40 years of age, have been reported globally, with the lowest rates found



© 2019 The Authors. JGH Open: An open access journal of gastroenterology and hepatology published by Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

23979070, 2020, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jgh3.12248 by South African Medical Research, Wiley Online Library on [06/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

23979070, 2020, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jgh3.12248 by South African Medical Research, Wiley Online Library on [06/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

in the United States, Europe, and New Zealand (1–6%) and the highest in Asian and Middle Eastern countries (10–39%).[11–13] The current SA National Cancer Registry (NCR) statistics, spanning a 5-year period (2010–2014), show that young CRC patients, under 50 years of age, contribute an average of 18% of all CRC cases, of which approximately 7% were younger than 40 years of age.[2–6] The NCR 2014 report also showed an exponential growth in CRC incidence in male patients from 2010 to 2014. A correlation between race and age was also demonstrated as the majority of young patients younger than 50 years old were black (32%) *versus* white patients (11%) ($P < 0.0001$) (NCR 2013–2014).[5,6] The burden of disease in SA has also been a challenge to assess due to underreporting of CRC cases as registries are mainly limited to small urban areas.[14]

Thus, this descriptive epidemiological study will serve as a preliminary study to obtain a greater understanding of the development of CRC, particularly in black SA CRC patients. This will include screening demographic and histopathological data in the hope of identifying key features associated with the disease to assist with earlier detection and improvement of prognosis, particularly within this population group.

## Methods

A retrospective laboratory information system (LIS) search was conducted on all patients who had biopsy samples or colorectal resections reported by the Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) branch of the National Health Laboratory Service (NHLS)/Anatomical Pathology Division, Faculty of Health Sciences, University of the Witwatersrand. Histological reports of all patients diagnosed with colorectal adenocarcinoma between 2011 and 2015 were reviewed, and data were described with respect to age, race, gender, tumor site ([distal to the splenic flexure [left-sided] and proximal to the splenic flexure [right-sided]), histological subtype and grade, presence and grade of precursor lesion, invasive tumor including the presence of tumor infiltrating lymphocytes (TIL), Crohn's-like inflammatory reaction (CIR), and the tumor stage (TNM). The Ethics Committee of the University of the Witwatersrand (clearance number M120994) approved this study.

***Statistical analysis.*** Statistical associations between race groups were investigated using Student's *t*-test, Fisher's exact test, and Chi Square test analysis. Non-black race patients were included in a single group, termed the "Other Race" group (White, Colored/mixed ancestry, Indian) due to small case numbers in the colored and Indian groups in comparison to the black patient group. Race was compared with age, tumor site, and morphological features as stated above. Statistical data analysis was completed using Stata Intercooled 7.0 (Stata, College Station, TX, USA) and Graphpad Prism version 7 (GraphPad Software, La Jolla, CA, USA). The results were considered statistically significant when $P < 0.05$* and highly significant when $P < 0.001$***.

## Results

A total of 665 CRC patients were diagnosed at CMJAH between 2011 and 2015, with the majority of patients being black (391/665; 59%) (Table 1). Overall, the patients were predominantly male (366/665; 55%), and no variation of gender was seen among race groups. The median age of CRC was 59 years (interquartile range: 49–68). A significant association was seen between patient age and race (black *vs* other race group [ORG]) (median age: 56 *vs* 62 respectively), ($P < 0.0001$). Most tumors occurred in the left colon (427/621; 69%), had an invasive adenocarcinoma subtype (596/646; 92%), were low-grade (LG) (568/624; 91%), and presented as TNM stage III (180/350; 52%). The majority of CRC cases did not contain TILs (224/345; 65%) or a CIR (263/345; 76%), nor did they show lymphatic invasion (296/422; 70%). A minority of polyps was found (140/408; 34%), with tubular adenomas (TAs) being reported as the predominant subtype (72/140; 51%), followed by tubulovillous adenomas (TVAs) (61/140; 44%). TAs mostly displayed LG dysplasia (40/67; 60%); however, TVAs were mostly associated with high-grade (HG) dysplasia (49/60; 82%) ($P < 0.0001$). All the aforementioned pathological characteristics lacked any correlation with race. BRAF V600E mutations, however, exclusively occurred in ORG patients (5/21; 24%) ($P = 0.0011$).

Correlation between age groups, young *versus* old (≤50 *vs* >50 years) and CRC subtype, tumor grade, TIL, and lymphatic invasion was identified (Table 2). Signet ring cell carcinoma occurred at a higher frequency in the young patient group (16/183; 9%) *versus* the older patient group (4/463; 1%) ($P = 0.0030$), while a mild to moderate TIL response was more commonly found in the older patient group (95/247; 38%) than in the younger group (26/98; 27%) ($P = 0.0450$). Lymphatic invasion was associated more with younger patients (45/120; 37%) than with the older group (81/302; 27%) ($P = 0.0341$), and BRAF V600E mutations were exclusively found in older patients (5/36; 24%) *versus* younger patients (0/40; 0%) ($P = 0.0204$). No significant association was found between gender and histopathological data (Table 3).

The majority of polyps were found in the left colon (81/131; 62%) and in male patients (89/140; 64%) (Tables 3 and 4). A particular association between the side of colon and grade of TVAs was noted here ($P = 0.0325$) as left-sided CRCs were mostly associated with HG TVAs (31/34; 91%). BRAF V600E mutations occurred exclusively in the right colon (5/34; 15%), and none were associated with left CRC tumors (0/42; 0%) ($P = 0.0151$).

## Discussion

The focus of this study was to ascertain common features, particularly demographic and histopathological characteristics, associated with black patients to assist in the earlier detection and better management of the disease. Although this study demonstrated no significant differences between male and female patients, black *versus* ORG and younger *versus* older patient groups illustrated significant differences in a number of parameters.

A total of 665 CRC patients were diagnosed at CMJAH over a 5-year period, from 2011 to 2015. Patients were predominantly male and black, with a considerable proportion being young. The median age of black patients compared to ORG patients (56 *vs* 62 years, respectively) was significantly younger. The data confirmed local and international reports


© 2019 The Authors. *JGH Open: An open access journal of gastroenterology and hepatology* published by Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

**Table 1** Descriptive histopathological analysis of colorectal cancer cases diagnosed at Charlotte Maxeke Johannesburg Academic Hospital between 2011 and 2015: Black *versus* other race groups

| Demographic data | Number of cases (%) | Stratified by race groups no. of cases (%) | | Statistical analysis: |
|---|---|---|---|---|
| | | Black (B) | Other race groups (O) | |
| Gender | 665 | 391 | 274 | $P = 0.9369$ |
| Male | 366 (55) | 216 (55) | 150 (55) | |
| Female | 299 (45) | 175 (45) | 124 (45) | |
| Age | 660 | 389 | 271 | ***P* < 0.0001*** |
| Median | 59 | 56*** | 62 | |
| Min–Max | 15–94 | 15–94 | 25–92 | |
| Mean ± SD | 57 ± 14 | 55 ± 15 | 62 ± 13 | |
| P25-P75 | 49–68 | 45–65 | 55–71 | |
| 95% CI | [56–59] | [53–56] | [60–64] | |
| ≤50 years | 184 (28) | 142 (37)*** | 42 (15) | ***P* < 0.0001*** |
| >50 years | 476 (72) | 247 (63) | 229 (85) | |
| Tumor site | 621 | 368 | 253 | $P = 0.7634$ |
| Left | 427 (69) | 250 (68) | 177 (70) | |
| Right | 190 (30) | 115 (31) | 75 (30) | |
| Left and right | 4 (1) | 3 (1) | 1 (0) | |
| Tumor subtype | 646 | 380 | 266 | $P = 0.2878$ |
| Invasive adenocarcinoma | 596 (92) | 349 (92) | 247 (93) | |
| Mucinous adenocarcinoma | 29 (5) | 16 (4) | 13 (5) | |
| Signet ring cell adenocarcinoma | 20 (3) | 15 (4) | 5 (2) | |
| Neuroendocrine carcinoma | 1 (0) | 0 (0) | 1 (0) | |
| Tumor grade | 624 | 372 | 252 | $P = 0.4793$ |
| LG | 568 (91) | 336 (90) | 232 (92) | |
| HG | 56 (9) | 36 (10) | 20 (8) | |
| AJCC TNM staging | 350 | 208 | 142 | $P = 0.4625$ |
| I | 46 (13) | 26 (13) | 20 (15) | |
| II | 109 (31) | 59 (28) | 50 (35) | |
| III | 180 (52) | 114 (55) | 66 (46) | |
| IV | 15 (4) | 9 (4) | 6 (4) | |
| Tumor-infiltrating lymphocytes | 345 | 202 | 143 | $P = 0.4231$ |
| None | 224 (65) | 135 (67) | 89 (62) | |
| Mild–moderate | 121 (35) | 67 (33) | 54 (38) | |
| Crohn's-like inflamatory response | 345 | 202 | 143 | $P = 0.3077$ |
| None | 263 (76) | 158 (78) | 105 (73) | |
| Mild–moderate | 82 (24) | 44 (22) | 38 (27) | |
| Lymphatic invasion | 422 | 254 | 168 | $P = 0.1935$ |
| Absent | 296 (70) | 172 (68) | 124 (74) | |
| Present | 126 (30) | 82 (32) | 44 (26) | |
| Polyps | 408 | 234 | 174 | $P = 0.2934$ |
| Absent | 268 (66) | 160 (68) | 108 (62) | |
| Present | 140 (34) | 74 (32) | 66 (38) | |
| Polyp subtype | 140 | 75 | 65 | $P = 0.1856$ |
| Hyperplastic polyp | 4 (3) | 1 (1) | 3 (5) | |
| Pseudopolyp | 2 (1) | 2 (3) | 0 (0) | |
| Sessile serrated adenoma | 1 (1) | 0 (0) | 1 (1) | |
| TA | 72 (51) | 43 (57) | 29 (45) | |
| TVA | 61 (44) | 29 (39) | 32 (49) | |
| TA grade | 67 | 39 | 28 | $P = 0.2161$ |
| LG | 40 (60) | 25 (65) | 15 (54) | |
| HG | 27 (40) | 14 (35) | 13 (46) | |
| TVA grade | 60 | 28 | 32 | |
| LG | 11 (18) | 2 (7) | 8 (25) | |
| HG | 49 (82) | 25 (93)*** | 24 (75) | |
| BRAF V600E | 76 | 55 | 21 | ***P* = 0.0011*** |
| Wild-type | 71 | 55 (100) | 16 (76) | |
| Mutation | 5 | 0 (0) | 5 (24)*** | |

*Statistically significant.

***Statistically highly significant.

AJCC, American Joint Committee on Cancer; CI, confidence interval; HG, high grade; LG, low grade; TA, tubular adenoma; TNM, tumor node metastases; TVA, tubulovillous adenoma.



© 2019 The Authors. *JGH Open: An open access journal of gastroenterology and hepatology* published by Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

23979070, 2020, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jgh3.12248 by South African Medical Research, Wiley Online Library on [06/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**Table 2** Descriptive histopathological analysis of colorectal cancer cases diagnosed at Charlotte Maxeke Johannesburg Academic Hospital between 2011 and 2015: Younger (≤50 years) *versus* older (>50 years) patients

| Demographic data | Number of cases (%) | Stratified by age groups no. of cases (%) | | Statistical analysis: |
|---|---|---|---|---|
| | | ≤50 years | >50 years | |
| Gender | 665 | 189 | 476 | $P = 0.1005$ |
| Male | 366 (55) | 114 (60) | 252 (53) | |
| Female | 299 (45) | 75 (40) | 224 (47) | |
| Age | 660 | 389 | 271 | ***P* < 0.0001*** |
| Median | 59 | 41 | 64 | |
| Min–Max | 15–94 | 15–50 | 51–94 | |
| Mean ± SD | 57 ± 14 | 40 ± 8 | 65 ± 9 | |
| P25-P75 | 49–68 | 34–46 | 57–71 | |
| 95% CI | [56–59] | [39–41] | [64–66] | |
| Tumor site | 621 | 179 | 442 | $P = 0.2999$ |
| Left | 426 (68) | 119 (66) | 307 (69) | |
| Right | 191 (31) | 60 (34) | 131 (30) | |
| Left and right | 4 (1) | 0 (0) | 4 (1) | |
| Tumor subtype | 646 | 183 | 463 | ***P* < 0.0001*** |
| Invasive adenocarcinoma | 596 (92) | 159 (87) | 437 (94) | |
| Mucinous adenocarcinoma | 29 (5) | 8 (4) | 21 (5) | |
| Signet ring cell adenocarcinoma | 20 (3) | 16 (9)*** | 4 (1) | |
| Neuroendocrine carcinoma | 1 (0) | 0 (0) | 1 (0) | |
| Tumor grade | 624 | 178 | 446 | ***P* = 0.0030**** |
| LG | 568 (91) | 152 (85) | 416 (93) | |
| HG | 56 (9) | 26 (15)** | 30 (7) | |
| AJCC TNM staging | 350 | 100 | 250 | $P = 0.2282$ |
| I | 46 (13) | 9 (9) | 37 (15) | |
| II | 109 (31) | 27 (27) | 82 (33) | |
| III | 180 (52) | 59 (59) | 121 (48) | |
| IV | 15 (4) | 5 (5) | 10 (4) | |
| Tumor-infiltrating lymphocytes | 345 | 98 | 247 | ***P* = 0.0450*** |
| None | 224 (65) | 72 (73) | 152 (62) | |
| Mild–moderate | 121 (35) | 26 (27) | 95 (38)* | |
| Crohn's-like inflamatory response | 345 | 98 | 247 | $P = 0.1613$ |
| None | 263 (76) | 80 (82) | 183 (74) | |
| Mild–moderate | 82 (24) | 18 (18) | 64 (26) | |
| Lymphatic invasion | 422 | 120 | 302 | ***P* = 0.0341*** |
| Absent | 296 (70) | 75 (63) | 221 (73) | |
| Present | 126 (30) | 45 (37)* | 81 (27) | |
| Polyps | 408 | 111 | 297 | $P = 0.1610$ |
| Absent | 268 (66) | 79 (71) | 189 (64) | |
| Present | 140 (34) | 32 (29) | 108 (36) | |
| Polyp subtype | 140 | 32 | 108 | $P = 0.6291$ |
| Hyperplastic polyp | 4 (3) | 0 (0) | 4 (4) | |
| Pseudopolyp | 2 (1) | 1 (3) | 1 (1) | |
| Sessile serrated adenoma | 1 (1) | 0 (0) | 1 (1) | |
| TA | 72 (51) | 18 (56) | 54 (50) | |
| TVA | 61 (44) | 13 (41) | 48 (44) | |
| Polyp grade | | | | $P = 0.6828$ |
| TA | 67 | 18 | 49 | |
| LG | 40 (60) | 10 (56) | 30 (60) | |
| HG | 27 (40) | 8 (44) | 19 (40) | |
| TVA | 60 | 12 | 48 | |
| LG | 11 (18) | 3 (25) | 8 (17) | |
| HG | 49 (82) | 9 (75) | 40 (83) | |
| BRAF V600E | 76 | 40 | 36 | ***P* = 0.0204*** |
| Wildtype | 71 | 40 (100) | 31 (76) | |
| Mutation | 5 | 0 (0) | 5 (24)* | |

*Statistically significant.

***Statistically highly significant.

AJCC, American Joint Committee on Cancer; CI, confidence interval; HG, high grade; LG, low grade; TA, tubular adenoma; TNM, tumor node metastases; TVA, tubulovillous adenoma.



© 2019 The Authors. *JGH Open: An open access journal of gastroenterology and hepatology* published by Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

23979070, 2020, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jgh3.12248 by South African Medical Research, Wiley Online Library on [06/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**Table 3** Descriptive histopathological analysis of colorectal cancer cases diagnosed at Charlotte Maxeke Johannesburg Academic Hospital between 2011 and 2015: Male *versus* female

| Demographic data | Number of cases (%) | Stratified by gender no. of cases (%) | | Statistical analysis: |
|---|---|---|---|---|
| | | Male | Female | |
| Gender | 665 | 366 (55%) | 299 (45%) | |
| Age | 660 | 364 | 296 | $P = 0.2560$ |
| Median | 59 | 58 | 60 | |
| Min–Max | 15–94 | 15–94 | 20–92 | |
| Mean ± SD | 57 ± 14 | 56 ± 14 | 59 ± 14 | |
| P25-P75 | 49–68 | 47–67 | 51–69 | |
| 95% CI | [56–59] | [55–58] | [58–61] | |
| Tumor site | 621 | 344 | 277 | $P = 0.9653$ |
| Left | 426 (68) | 237 (69) | 189 (68) | |
| Right | 191 (31) | 105 (31) | 86 (31) | |
| Left and right | 4 (1) | 2 (0) | 2 (1) | |
| Tumor subtype | 646 | 183 | 463 | $P = 01663$ |
| Invasive adenocarcinoma | 596 (92) | 437 (87) | 267 (94) | |
| Mucinous adenocarcinoma | 29 (5) | 21 (4) | 14 (5) | |
| Signet ring cell adenocarcinoma | 20 (3) | 4 (9) | 9 (1) | |
| Neuroendocrine carcinoma | 1 (0) | 1 (0) | 0 (0) | |
| Tumor grade | 624 | 344 | 280 | $P = 01611$ |
| LG | 568 (91) | 308 (85) | 260 (93) | |
| HG | 56 (9) | 36 (15) | 20 (7) | |
| AJCC TNM staging | 350 | 100 | 250 | $P = 0.2282$ |
| I | 46 (13) | 9 (9) | 37 (15) | |
| II | 109 (31) | 27 (27) | 82 (33) | |
| III | 180 (52) | 59 (59) | 121 (48) | |
| IV | 15 (4) | 5 (5) | 10 (4) | |
| Tumor-infiltrating lymphocytes | 345 | 192 | 153 | $P = 0.1119$ |
| None | 224 (65) | 132 (69) | 92 (60) | |
| Mild–moderate | 121 (35) | 60 (31) | 61 (40) | |
| Crohn's-like inflamatory response | 345 | 192 | 153 | $P = 0.0569$ |
| None | 263 (76) | 154 (80) | 109 (71) | |
| Mild–moderate | 82 (24) | 38 (20) | 44 (29) | |
| Lymphatic invasion | 422 | 234 | 188 | $P = 0.6698$ |
| Absent | 296 (70) | 162 (69) | 134 (71) | |
| Present | 126 (30) | 72 (31) | 54 (29) | |
| Polyps | 408 | 236 | 174 | $P = 0.0917$ |
| Absent | 268 (66) | 147 (62) | 123 (71) | |
| Present | 140 (34) | 89 (38) | 51 (29) | |
| Polyp subtype | 140 | 89 | 51 | $P = 0.4160$ |
| Hyperplastic polyp | 4 (3) | 3 (4) | 1 (2) | |
| Pseudopolyp | 2 (1) | 2 (2) | 0 (0) | |
| Sessile serrated adenoma | 1 (1) | 0 (0) | 1 (2) | |
| TA | 72 (51) | 43 (48) | 29 (57) | |
| TVA | 61 (44) | 41 (46) | 20 (39) | |
| Polyp grade | | | | $P = 0.4732$ |
| TA | 67 | 36 | 28 | |
| LG | 40 (60) | 21 (58) | 17 (54) | |
| HG | 27 (40) | 15 (42) | 11 (46) | |
| TVA | 60 | 41 | 19 | |
| LG | 11 (18) | 6 (15) | 5 (28) | |
| HG | 49 (82) | 35 (85) | 14 (72) | |
| BRAF V600E | 76 | 44 | 32 | $P = 0.6444$ |
| Wildtype | 71 | 42 (95) | 29 (91) | |
| Mutation | 5 | 2 (5) | 3 (9) | |

*Statistically significant.

***Statistically highly significant.

AJCC, American Joint Committee on Cancer; CI, confidence interval; HG, high grade; LG, low grade; TA, tubular adenoma; TNM, tumor node metastases; TVA, tubulovillous adenoma.



© 2019 The Authors. *JGH Open: An open access journal of gastroenterology and hepatology* published by Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

23979070, 2020, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jgh3.12248 by South African Medical Research, Wiley Online Library on [06/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**Table 4** Descriptive histopathological analysis of CRC cases diagnosed at Charlotte Maxeke Johannesburg Academic Hospital between 2011–2015: Left *versus* right-sided colon cancer

| Demographic data | Number of cases (%) | Stratified by tumor site no. of cases (%) | | Statistical analysis: |
|---|---|---|---|---|
| | | Left-sided | Right-sided | |
| Tumor site | 617 | 426 (69%) | 191 (31%) | |
| Tumor subtype | 607 | 419 | 188 | ***P* < 0.0001*** |
| Invasive adenocarcinoma | 596 (92) | 398 (95) | 162 (86) | |
| Mucinous adenocarcinoma | 29 (5) | 9 (2) | 18 (10)*** | |
| Signet ring cell adenocarcinoma | 20 (3) | 12 (3) | 7 (4) | |
| Neuroendocrine carcinoma | 1 (0) | 0 (0) | 1 (0) | |
| Tumor grade | 587 | 399 | 188 | *P* = 02845 |
| LG | 533 (91) | 366 (92) | 167 (89) | |
| HG | 54 (9) | 33 (8) | 21 (11) | |
| AJCC TNM staging | 350 | 192 | 153 | *P* = 0.6086 |
| I | 46 (13) | 28 (15) | 17 (11) | |
| II | 109 (31) | 60 (31) | 46 (30) | |
| III | 180 (52) | 96 (50) | 80 (52) | |
| IV | 15 (4) | 8 (4) | 10 (7) | |
| Tumor-infiltrating lymphocytes | 338 | 193 | 145 | ***P* = 0.0210*** |
| None | 220 (65) | 136 (70) | 84 (58) | |
| Mild–moderate | 118 (35) | 57 (30) | 61 (42)* | |
| Crohn's like inflamatory response | 338 | 193 | 145 | *P* = 1.0000 |
| None | 258 (76) | 147 (76) | 111 (77) | |
| Mild–moderate | 80 (24) | 46 (24) | 34 (23) | |
| Lymphatic invasion | 417 | 254 | 163 | ***P* = 0.0009**** |
| Absent | 293 (70) | 194 (76) | 99 (61) | |
| Present | 123 (30) | 60 (24) | 64 (39)*** | |
| Polyps | 392 | 231 | 161 | *P* = 0.4468 |
| Absent | 261 (67) | 150 (65) | 111 (69) | |
| Present | 131 (33) | 81 (35) | 50 (31) | |
| Polyp subtype | 131 | 81 | 50 | *P* = 0.1484 |
| Hyperplastic polyp | 3 (2) | 3 (4) | 0 (0) | |
| Pseudopolyp | 2 (1) | 0 (0) | 2 (4) | |
| Sessile serrated adenoma | 1 (1) | 0 (0) | 1 (2) | |
| TA | 68 (52) | 43 (53) | 25 (50) | |
| TVA | 57 (44) | 35 (43) | 22 (44) | |
| Polyp grade | 119 | | | *P* = 0.0615 |
| TA | 64 | 40 | 24 | |
| LG | 38 (59) | 21 (52) | 17 (71) | |
| HG | 26 (41) | 19 (48) | 7 (29) | |
| TVA | 55 | 34 | 21 | |
| LG | 10 (18) | 3 (9) | 7 (33) | |
| HG | 45 (82) | 31 (91)* | 14 (67) | |
| BRAF V600E | 76 | 42 | 34 | ***P* = 0.0154*** |
| Wildtype | 71 | 42 (100) | 29 (85) | |
| Mutation | 5 | 0 (0) | 5 (15)* | |

*Statistically significant.

***Statistically highly significant.

AJCC, American Joint Committee on Cancer; CI, confidence interval; CRC, colorectal cancer; HG, high grade; LG, low grade; TA, tubular adenoma; TNM, tumor node metastases; TVA, tubulovillous adenoma.

regarding male *versus* female ratios (55 *vs* 45%, respectively),[1,5] with black patients showing an earlier age of onset in comparison to the white group.[7,15] In comparison, in African Americans (AA), the median age of onset in males and females were reported to be 66 and 70 years, respectively, compared to 72 and 77 in white men and women, respectively.[16] In this SA cohort, the median age for male and female black patients was 55 and 56 years, respectively, *versus* 62 and 65 years for males and females of ORG, respectively. The median age for white male and female patients (separated from Indian and colored or mixed race) remained 62 and 65 years, respectively. These data showed that SA black and white patients are at least 10 years younger in comparison to U.S. AA and U.S. white patients.



© 2019 The Authors. *JGH Open: An open access journal of gastroenterology and hepatology* published by Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

23979070, 2020, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jgh3.12248 by South African Medical Research, Wiley Online Library on [06/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The frequency of young black patients younger than 50 years of age in this cohort (37%) was similar to that of young black SA CRC patients in the NCR 2014 report (32%). The majority of tumors occurred in the left colon, had an invasive adenocarcinoma subtype, LG tumors, and presented at an advanced TNM stage (III and IV) These findings were in concordance with international published data that reported an approximate 70% occurrence of left-sided CRC, with about 90% being of an invasive adenocarcinoma subtype, that was moderately to well differentiated, with approximately 50–60% of CRC cases presenting at a more advanced stage.[17,18] Male and female patients had equal frequencies for tumor site (69% left-sided *vs* 31% right-sided). This finding is contrary to other studies showing right-sided colon cancer to be increased in female patients and left-sided CRC to be increased in males.[19]

In the present analysis, with regard to polyps, conventional adenomas (TA, TVA, and villous adenomas [VAs]) are mostly found in the left colon,[20] with TAs described as the predominant subtype (80%), TVAs accounting for 10–20%, and VAs less than 5%.[21–23] In this study, polyps were found in 34% (140/408) of cases, with only 1% being sessile serrated adenoma (SSA), 4% hyperplastic polyps (HPs), 51% TAs, and 44% TVAs. TAs were mostly associated with LG dysplasia (60%), and TVAs that also occurred at a considerably high frequency than that stated in literature were mostly associated with HG dysplasia (83%) ($P < 0.0001$). The size of the adenomatous polyp (AP) was not recorded in the majority of reports, only the quantity of APs and grade of dysplasia. This study showed no association of APs with race, gender, or patient age; however, HG TVAs were more commonly found in the left colon ($P = 0.0325$). The transition from LG APs to more advanced APs or CRC takes approximately 3 years, suggesting that frequent surveillance colonoscopies for the earlier detection of CRC would benefit patients. Guidelines indicate repeat surveillance every 5–10 years for 1–2 small TAs (<1 cm) with LG dysplasia, and where 3–10 large APs are found (≥1 cm) or in those with HG dysplasia or with villous features, colonoscopy should be performed every 3 years.[20,23]

The literature indicates that AA patients are more likely to have right-sided tumors than left.[15,24,25] This SA study, however, showed no relationship between race and tumor site. BRAF V600E mutational analysis was carried out in 76 patients, and an association was only found in older white patients with right-sided colon cancers, possibly indicating development via the microsatellite instability CRC pathway.

CRC research in AA patients has demonstrated higher incidence and mortality rates, earlier age of diagnosis, more advanced stage of the disease, and a greater proportion of right-sided colon cancers in comparison to other ethnic groups. This has led American professional societies to change guidelines, recommending colonoscopy screening by 45 instead of 50 years of age, which has led to a great reduction in CRC incidence in this population group.[26] Although statistically significant differences concerning stage and tumor site were not seen when comparing black *versus* ORG patients, later stage of disease and earlier age of onset in black patients remained significant factors, and a policy recommending earlier screening in this population group should be considered to reduce incidence rates.

In conclusion, this study shows that black SA CRC patients most likely follow a different carcinogenic pathway in comparison to other SA ethnic groups due to earlier age of onset. Black SA patients also seem to present at least a decade earlier when comparing Africans found in the Western world, with more left-sided CRC occurring in SA patients compared to a right-sided CRC association seen in AAs. Moreover, a lack of BRAF V600E mutation in the black SA CRC population in this study suggests that there is a different carcinogenic pathway in black SA CRC patients compared to other ethnic groups, which still needs to be evaluated in a larger cohort of patients. TVAs found at a frequent rate across different ethnic groups in SA patients could possibly suggest a dominant adenocarcinoma sequence in the development of the disease. The initiation of earlier screening and more frequent surveillance policies in black SA patients could possibly result in better management of the disease. Further epidemiological research linked to molecular characterization in SA population groups is still needed to assist in determining the underlying cause and designing more personalized treatment strategies for better management of the disease.

## Acknowledgments

The authors acknowledge Dr P. Magangane for statistical analysis contribution. This work was funded by Griffin Trust Fund (GTCP016), National Health Laboratory Services Trust Fund (94347) and Wits University FRC Fund (001.251. 8462101.5121105.000000.0000000000.4550).

## References

1 Ferlay J, Soerjomataram I, Dikshit R *et al*. Cancer incidence and mortality - major patterns in GLOBOCAN 2012, worldwide and Georgia. *Int. J. Cancer*. 2015; **136**: E359–86.
2 South African National Cancer Registry. *Cancer in South Africa (2010)*. Cited Oct 25 2018. Available from URL: http://www.cansa.org.za/south-african-cancer-statistics
3 South African National Cancer Registry. *Cancer in South Africa (2011)*. Cited Oct 25 2018. Available from URL: http://www.cansa.org.za/south-african-cancer-statistics/
4 South African National Cancer Registry. *Cancer in South Africa (2012)*. Cited Oct 25 2018. Available from URL: http://www.cansa.org.za/south-african-cancer-statistics/
5 South African National Cancer Registry. *Cancer in South Africa (2013)*. Cited Oct 25 2018. Available from URL: http://www.cansa.org.za/south-african-cancer-statistics/
6 South African National Cancer Registry. *Cancer in South Africa (2014)*. Cited Oct 25 2018. Available from URL: www.ncr.ac.za
7 Cronjé L, Paterson AC, Becker PJ. Colorectal cancer in South Africa: a heritable cause suspected in many young black patients. *S. Afr. Med. J.* 2009; **99**: 103–6.
8 Katsidzira L, Gangaidzo I, Thomson S, Rusakaniko S, Matenga J, Ramesar R. The shifting epidemiology of colorectal cancer in sub-Saharan Africa. *Lancet Gastroenterol. Hepatol.* 2017; **2**: 377–83.
9 Bebington B. An evaluation of epidemiological differences in colorectal cancer patients over and under the age of 50 presenting to academic centers in Johannesburg. *S. Afr. J. Surg.* 2017; **55**: 54.
10 Prodehl L, Bebington B, Fabian J, Singh E, Ruff P. Colorectal cancer in a South Africa urban setting - a preliminary analysis. *S. Afr. J. Surg.* 2017; **55**: 58.


© 2019 The Authors. *JGH Open: An open access journal of gastroenterology and hepatology* published by Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

23979070, 2020, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jgh3.12248 by South African Medical Research, Wiley Online Library on [06/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

11 Gupta S, Acharya AN. Colorectal carcinoma in the young, colorectal cancer - from prevention to patient care. In: Ettarh R (ed.). *Intechopen*. InTech Europe, 2012; 538. Available from URL: http://www.intechopen.com/books/colorectal-cancer-from-prevention-to-patient-care/colorectal-carcinoma-inthe-young.
12 Campos FG. Colorectal cancer in young adults: a difficult challenge. *World J. Gastroenterol.* 2017; **23**: 5041–4.
13 Bouassida M, Feidi B, Mroua B *et al*. Histopathologic characteristics and short-term outcomes of colorectal cancer in young tunisian patients: one center's experience. *Pan Afr. Med. J.* 2012; **12**: 1–7.
14 Bebington B, Singh E, Fabian J *et al.* Design and methodology of a study on colorectal cancer in Johannesburg, South Africa. *JGH Open.* 2018; **2**: 139–43.
15 Williams R, White P, Nieto J, Vieira D, Francois F, Hamilton F. Colorectal Cancer in African Americans: an Update. *Clin. Transl. Gastroenterol.* 2016; **7**: e185.
16 DeSantis CE, Siegel RL, Sauer AG *et al*. Cancer statistics for African Americans, 2016: progress and opportunities in reducing racial disparities. *CA Cancer J. Clin.* 2016; **66**: 290–308.
17 Yamagishi H, Kuroda H, Imai Y, Hiraishi H. Molecular pathogenesis of sporadic colorectal cancers. *Chin. J. Cancer*. 2016; **35**: 1–8.
18 Ansa B, Coughlin S, Alema-Mensah E, Smith S. Evaluation of colorectal cancer incidence trends in the United States (2000–2014). *J. Clin. Med.* 2018; **7**: 22.
19 Brenner H, Hoffmeister M, Arndt V, Haug U. Gender differences in colorectal cancer: implications for age at initiation of screening. *Br. J. Cancer*. 2007; **96**: 828–31.
20 Munding J, Tannapfel A. Epidemiology of colorectal adenomas and histopathological assessment of endoscopic specimens in the colorectum. *Viszeralmedizin*. 2014; **30**: 10–6.
21 Marques S, Rosa VF, Carlos A. Influence of patient age and colorectal polyp size on histopathology findings. *Arq. Bras. Cir. Dig.* 2014; **27**: 109–13.
22 Park SJ, Yoon H, Jung IS *et al*. Clinical outcomes of surveillance colonoscopy for patients with sessile serrated adenoma. *Intest. Res.* 2018; **16**: 134–41.
23 Myers D, Arora K. Adenoma, villous. In: *StatPearls*. Treasure Island, Florida StatPearls Publ, 2019; 4–7. Cited May 23 2019. Available from URL:. https://www.ncbi.nlm.nih.gov/books/NBK470272/.
24 Shavers VL. Racial/ethnic variation in the anatomic subsite location of in situ and invasive cancers of the colon. *J. Natl. Med. Assoc.* 2007; **99**: 733–48.
25 Dimou A, Syrigos KN, Saif MW. Disparities in colorectal cancer in African-Americans vs Whites: before and after diagnosis. *World J. Gastroenterol.* 2009; **15**: 3734–43.
26 DeSantis C, Naishadham D, Jemal A. Cancer statistics for African Americans. *CA Cancer J. Clin.* 2013; **63**: 151–66.



© 2019 The Authors. *JGH Open: An open access journal of gastroenterology and hepatology* published by Journal of Gastroenterology and Hepatology Foundation and John Wiley & Sons Australia, Ltd.

23979070, 2020, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jgh3.12248 by South African Medical Research, Wiley Online Library on [06/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

# Chapter 3:

Microsatellite Instability assessment in Black South African Colorectal Cancer patients reveal an increased incidence of suspected Lynch syndrome.

SCIENTIFIC REPORTS
nature research

OPEN

# Microsatellite Instability assessment in Black South African Colorectal Cancer patients reveal an increased incidence of suspected Lynch syndrome

M. McCabe[1*], Y. Perner[1], R. Magobo[1], P. Magangane[1], S. Mirza[2] & C. Penny[2]

**Microsatellite Instability (MSI) is a hallmark of colorectal cancer (CRC) and occurs in 15–16% of CRC. Molecular biological information of CRC in South Africa (SA) is largely unrecorded. This study was undertaken to determine the frequency of MSI, with particular reference to Lynch syndrome (LS) with a view to improve surveillance and prevention strategies. This was a retrospective study on CRC samples diagnosed between 2011–2015 at Charlotte Maxeke Johannesburg Academic Hospital (CMJAH). Samples diagnosed between 2011–2012 were screened for MSI by PCR and mismatch repair (MMR) immunohistochemistry (IHC), and additional BRAFV600E mutational analysis performed. T-tests, Fischer's exact and Chi square statistical tests were applied. Twelve percent of patients displayed MSI, with increased frequency in black (15%) versus other ethnic group (OEG) (8%) patients. MSI patients were significantly younger than microsatellite stable (MSS) patients, however when stratified by ethnicity, black patients were predominantly younger (median age: 47), with increased MSH2/6 loss, and no BRAF mutations. These findings suggest a large proportion of young black SA CRC patients develop via the LS pathway due to earlier age onset and predominant MSH2/6 protein loss. SA patients of other ethnicities appear to follow the more well established sporadic MSI pathway.**

Colorectal cancer (CRC) develops through three major molecular pathways namely: Chromosomal Instability (CIN), Microsatellite Instability (MSI), and the epigenetic instability or CpG island methylator phenotype (CIMP) pathway[1–3]. CIN or microsatellite stable (MSS) tumours account for 65–75% of CRC and mainly develop through mutant adenomatous polyposis coli (APC) gene with subsequent KRAS mutational activation, TP53 inactivation and somatic copy number alterations (SCNAs)[4,5]. MSI CRC occurs in approximately 15–16% of CRC due to inactivation of the DNA Mismatch Repair (MMR) system (MLH1, MSH2, MSH6, PMS2), with the majority (12%) presenting as sporadic MSI CRC as a result of epigenetic mutation (methylation silencing) of the promoter sequence of MLH1, and BRAF V600E oncogenic mutations[2,6–10]. The remaining subset (2–3%) is caused by germline mutations in either one of 4 MMR genes or the EpCAM gene and is a feature of Lynch syndrome (LS)[8,9,11–14]. CIMP CRC (10–20%) demonstrate hypermethylation of several promoter CpG island loci throughout the genome, leading to tumour suppressor and tumour-related gene inactivation. The CIMP and the MSI pathway overlaps in sporadic MSI CRCs, as these tumours display high levels of CIMP and MSI and are characterized by a different type of precursor lesion in comparison to LS[15,16]. LS follows the classic adenoma-carcinoma sequence pathway as they present mainly with tubular adenomas (TAs) or tubulovillous adenomas (TVAs), whereas sporadic MSI CRCpremalignant lesions are sessile serrated adenomas (SSAs) developing through the serrated neoplasia pathway[15–17]. Clinicopathological features of sporadic MSI CRC include predominant occurrence in female patients, within the right colon and associated morphology includes signet ring cell and mucinous features with

[1]Department of Anatomical Pathology, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Parktown, Johannesburg, 2193, South Africa. [2]Department of Internal Medicine, Faculty of Health Sciences, University of the Witwatersrand, Parktown, Johannesburg, 2193, South Africa. *email: michelle.mccabe@nhls.ac.za

tumour infiltrating lymphocytes (TIL)[15,18–20]. LS patients are mainly associated with younger patients, with no gender preference and similar morphology as sporadic MSI tumours[12].

Previous CRC studies conducted in SA showed a higher frequency of MSI CRC in young black patients through MMR deficiency than white patients[21–23]. These studies suggested a high frequency of LS, however additional validation studies and further molecular characterization of MSI CRC in black SA patients was recommended[24].

MSI assessment has also shown important prognostic and predictive roles in CRC patient care, as MSI tumours are associated with a better prognosis than MSS tumours in early stage CRC, and treatment of MSI stage II tumours are not well responsive to 5-fluorouracil (5-FU) (standard treatment)[25,26]. MSI/BRAF wild-type tumours are also more suggestive of LS and is crucial for improving cancer surveillance and prevention screening for patient family members, due to their increased risk of developing cancer. Identifying LS patients and treating with aspirin (600 mg per day) have also shown to reduce the risk of CRC[27].

This study assesses the frequency and features associated with MSI CRC over a 5-year period in a cohort of newly diagnosed CRC patients at the Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) within SA. This data will provide insight into the CRC histopathological and molecular features associated with MSI CRC in black SA patients, with particular reference to MSI CRC frequency and the occurrence of suspected LS, a heretofore largely unassessed aspect of the disease.

## Methodology

**Patient demographics and tumour pathological characterisation.** This was a retrospective study, comprising a 5-year cohort of 675 patients who had biopsy samples or colorectal resections fulfilling the histological criteria for adenocarcinoma of the colon or rectum from January 2011–December 2015, reported by the Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) branch of the National Health Laboratory Service (NHLS)/Anatomical Pathology Division, Faculty of Health Sciences, University of the Witwatersrand. A total of 439 CRC cases with an MSI status by MMR immunohistochemistry (IHC) or MSI PCR result were included in this study. All cases were stratified by age, gender, ethnicity, tumour site, histological subtype, grade, stage (TNM classification by AJCC staging), presence and grade of precursor lesion, presence of tumour infiltrating lymphocytes (TIL) and Crohn's-like inflammatory reaction (CIR). Tumour site was considered left-sided if distal from the splenic flexure and right-sided if proximal to the splenic flexure. This information was obtained from histology reports. A semi-quantitive H&E assessment using the Klintrup-Mäkinen score was used to determine the presence of TILs[28]. Immunophenotypic profiling of TILs was not performed. Family histories were not available from the histological reports. This research was conducted under the institution's blanket ethics clearance (M10744) which allows for research to be carried out on all archived pathology specimens without informed consent from study participants. Additional study-specific ethics clearance was also obtained by the Ethics Committee of the University of the Witwatersrand (clearance number M120994), and all tests were performed according to the relevant guidelines and regulations.

**Molecular characterisation.** Molecular characterization of patient samples for frequency assessment was limited to the first 2 years (2011–2012) mainly due to the availability of funds (Figure 1). A total of 275 CRC cases were diagnosed during this period and a cohort of 267 archival CRC tumour specimens was retrieved, examined and subjected to molecular analysis. MSI PCR was conducted and samples with a MSI-High profile had additional BRAF V600E mutational analysis performed.

Cases without MMR IHC results as ascertained from the histology reports were screened for MMR protein loss by IHC.

DNA was extracted from 4 × 10μm sections of tumour tissue from formalin fixed paraffin embedded (FFPE) blocks using the QIAamp DNA FFPE tissue kit (Qiagen, Valencia, CA USA). The extracted DNA was subjected to a pentaplex PCR panel of 5 markers comprising mononucleotide repeats amplified by 5 sets of primers, i.e. (BAT-25, BAT-26, NR-21, NR-24, and NR-27) as described in Haghighi *et al.*[29]. Each antisense primer was labeled with 1 of the fluorescent markers FAM, HEX, or NED. The Applied Biosystems® True Allele® PCR master mix was used according to manufacturer's instructions with slight modifications for successful amplification of the multiplex PCR. Fluorescent PCR products were subjected to capillary electrophoresis on the Applied Biosystems PRISM 3500 automated genetic analyzer, and allelic size was determined using Genemapper 4.0 software (Applied Biosystems). Samples with no allelic size variations in the 5 microsatellite markers were classified as microsatellite stable (MSS), allelic size variations found in 1 of the 5 microsatellites markers were classified as microsatellite instability-low (MSI-L) and variation in 2 or more microsatellite markers was classified as microsatellite instability-high (MSI-H)[29]. Previous studies have indicated black people show polymorphic variation frequencies within alleles NR21, BAT25 and BAT26 of approximately 10%[30–32]. For tumour samples that demonstrated instability in only 2 of these 3 markers, matched normal tissue was assessed to determine the instability status. In cases where no normal tissue could be assessed, samples were regarded as MSS.

MMR IHC was performed on MSI-H samples that were not previously assessed, utilizing the Dako EnVision™ FLEX detection system Kit and Autostainer Link 48 equipment (Dako, Glostrup, Denmark) according to manufacturer's instructions. The primary monoclonal antibodies used were FLEX monoclonal mouse antibodies against MLH1 (clone ES05, ref ISO79, ready-to-use), MSH2 (clone FE11, ref IRO85, ready-to-use), MSH6 (clone EP49, ref IRO86, ready-to-use) and PMS2 (clone EP51, IRO87, ready-to-use) (Dako, Glostrup, Denmark). These were applied to 4-μm deparaffinized formalin fixed paraffin-embedded (FFPE) tissue sections. Blockage of endogenous peroxide activity was carried out with the EnVision™ FLEX Peroxidase-Blocking Reagent (Dako, Glostrup, Denmark) and Antigen retrieval performed at 97 °C for 20 minutes (pH 9.0). Antigen-antibody reaction was visualized with the DAB (Diaminobenzidine) solution, and subsequently counter-stained with hematoxylin. Each sample was screened by an expert pathologist for the absence or presence of expression of each of the

4 MMR proteins. Negative protein expression was defined as complete absence of nuclear staining within tumor cells. Expression of MMR proteins in normal epithelium, lymphocytes and stromal cells was used as positive internal controls for each case. The four deficient (d) MMR patterns assessed were 1: dMLH1/PMS2; 2: dMSH2/MSH6; 3: dMSH6 and 4: dPMS2.

All MSI-H samples were assessed for BRAF V600E mutations using the real-time Therascreen RGQ BRAF V600E mutation kit (Qiagen, Whitehead Scientific WHS), according to the manufacturer's instructions. The kit uses ARMS® and Scorpions® technology, and includes a control assay to evaluate the adequacy of quality and quantity of DNA in a sample, before assessing the presence or absence of mutated DNA. Results were only deemed valid if the control run passed.

Data obtained from the molecular characterization study (2011–2012) led to a more comprehensive study which included cases with an MSI status from histological reports for the total 5-year study period (2011–2015).

**Statistical analysis.** All clinical data was collected in an excel sheet and statistical analysis was performed using Stata Intercooled 7.0 (Stata, College Station, TX, USA) and Graphpad Prism version 7 (GraphPad Software, La Jolla, CA, USA). MSS and MSI-L cases were grouped together into one group (MSS), to establish unique associations with the counterpart MSI-H group. Patients were also analysed on the basis of ethnicity: black (African) versus other ethnic groups i.e., white (Caucasian), coloured (Mixed Ancestry), Indian (Asian); age and gender. The ability to draw statistically significant results on all ethnic groups was not possible due to small case numbers of certain groups. For this reason, White, Indian and Coloured patients were grouped into one "Other Ethnic Group" (OEG) to allow comparison with their black counterpart. Histopathological features were assessed as stated in the methodology section, and associations were evaluated using Fischer's exact and Chi-square tests. A result with a P value less than 0.05 was considered statistically significant.

## Results

**Patient demographics.** A total 439 of 672 CRC patients diagnosed between January 2011-December 2015 had MSI CRC status assessment. Of these cases, 247 (56%) were male and 192 (44%) female, with an overall median age of 58 years. Black patients contributed to 60% (262/439) of cases. (Column1 of Table 1).

**Pathological characteristics.** Further descriptive analysis was done with respect to histological characteristics, with left sided tumours accounting for the majority of cases (63%; 273/430) (Column1 of Table 1). The most common subtype was invasive adenocarcinoma (384/430; 89%), followed by invasive mucinous adenocarcinoma (30/430; 7%). AJCC tumour stage II and III accounted for the majority of cases (259/287; 90%) (Table 1). Further stratification by MSI and MSS subtype is illustrated in the 2nd and 3rd column in Table 1.

**MSI CRC molecular characterization (2011–2012).** MSI PCR performed on 267 cases (2011–2012) illustrated the majority were MSS (191/267; 71%), followed by MSI-L (45/267; 17%) and MSI-H (31/267; 12%). MMR IHC results demonstrated dMLH1 in most MSI CRC cases (13/31; 42%), followed by dMSH2/6 (10/31; 32%), dPMS2 (2/31; 6%) and dMSH6 (1/31; 3%). BRAF mutational analysis was performed on all MSI-H cases with only 10% (3/31) harbouring a mutation, and all 3 (3/13; 23%) occurred in dMLH1 tumours (See Fig. 1).

**MSI versus MSS CRC subtyping (2011–2015).** Male patients comprised 56% of the entire CRC cohort. Correlation between subtype and age was found ($P = 0.0033$) as median age was significantly lower in MSI CRC compared to the MSS subtype (51 vs 59 years respectively) (Table 1). A higher frequency of MSI tumours was seen in blacks ([22/148; 15%] versus OEG patients [9/119; 8%]), for the period 2011–2012. A significant relationship was seen between CRC subtype and tumour site ($P < 0.0001$), as MSS was associated with left sided tumours (261/370; 70%) and MSI with the right sided tumours [(47/60; 78%). A link between CRC molecular subtype and tumour subtype was also found ($P < 0.0001$), as mucinous adenocarcinoma occurred at a more frequent rate in MSI CRC; (13/60; 22%) than MSS CRC (17/370; 5%). Tumour infiltrating lymphocytes (TIL) response and subtype was found to have a correlation ($P = 0.0045$) as this occurred more often in the MSI group (26/47; 55%) than the MSS (77/236; 33%). Polyp subtypes tubular adenoma (TA) and tubulovillous adenoma (TVA) were associated with molecular subtypes ($P = 0.0216$) as TAs occurred more frequently in MSS CRC (49/84; 58%) and TVAs in MSI CRC (11/17; 65%).

Within the MSI CRC subgroup, additional association evaluations between ethnicity, gender, tumour site, defective MMR proteins and BRAFV600E mutation status were performed (Table 2). A correlation with age was shown ($P = 0.0006$) as black patients were significantly younger compared to OEG patients (median age: 47 versus 62, respectively). Tumour site and ethnicity showed an association for 2011–2012 data ($P = 0.0105$), as the majority of tumours were right sided within OEG patients (8/9; 89%), whereas black patients had left (8/22; 36%) and right-sided tumours (13/22; 59%) (Supplementary Data Table 2). Data for 2011–2015 however showed no such association ($P = 0.1817$). An association between ethnic groups and dMMR protein expression was found ($P = 0.0098$), where dMLH1/PMS2 expression was more commonly seen in OEG patients (15/22; 68%), and dMSH2/6 in black patients (19/39; 49%). BRAF V600E mutation status and ethnicity were related in the 2011–2012 data ($P = 0.0187$) (Supplementary Data Table 2), as all black patients were wild-type for BRAF gene status, however within the OEG, BRAF mutations were only found in dMLH1 tumours (3/5; 60%). For the 2011–2015 BRAFV600E data analysis, 2011–2012 results were dominantly represented with an additional 3 cases from 2013–2015 reports with BRAFV600E status. An increased association between BRAFV600E and OEG patients was shown ($P = 0.0045$) for this period.

Table 3 illustrates a multivariate analysis including variables considered clinically significant for MSI status. Total number of observations were 277. Age, tumour site, mucinous adenocarcinomas and TILS associated after controlling for confounding factors. Patients >50 years were 67% less likely to have MSI and right sided tumours were 8X more likely to have MSI. Invasive adenocarcinomas were 73% less likely to show MSI than mucinous

| | Number of cases (%) | CRC subtype | | Statistical analysis: Significance indicated by* |
|---|---|---|---|---|
| **Frequency:** | **All CRC cases** | **MSI** | **MSS** | |
| **2011–2015** | **439** | **61 (14)** | **378 (86)** | |
| **Demographical data (2011–2015)** | | | | |
| **GENDER** | **439** | **61** | **378** | |
| Male | 247 (56) | 34 (56) | 213 (56) | P = 1.0000 |
| Female | 192 (44) | 27 (44) | 165 (44) | |
| **AGE** | **436** | **61** | **375** | |
| Min-Max | 15–92 | 27–77 | 15–92 | |
| Mean ± SD | 56 ± 14 | 52 ± 13 | 57 ± 14 | **P = 0.0033**** |
| Median | 58 | 51** | 59 | |
| P25-P75 (Interquartile Range) | 47–67 | 40–62 | 48–68 | |
| 95% CI | [55–58] | [48–55] | [56–59] | |
| **Categorical Age** | **436** | **61** | **375** | |
| ≤50 years | 137 (31) | 29 (48)** | 108 (29) | |
| >50 years | 299 (69) | 32 (52) | 267 (71) | **P = 0.0046**** |
| **Ethnic groups** | **439** | **61** | **378** | |
| Black | 262 (60) | 39 (64) | 223 (59) | P = 0.4858 |
| Other Ethnic Groups (OEG) | 177 (40) | 22 (36) | 155 (41) | |
| **Histological characteristics (2011–2015)** | | | | |
| **TUMOUR SITE** | **430** | **60** | **370** | |
| Left (distal from splenic flexure) | 273 (63) | 12 (20) | 261 (70) | **P < 0.0001***** |
| Right (proximal to splenic flexure) | 154 (36) | 47 (78)*** | 107 (29) | |
| Left and Right | 3 (1) | 1 (2) | 2 (1) | |
| **TUMOUR SUBTYPE** | **430** | **60** | **370** | |
| Invasive Adenocarcinoma | 384 (89) | 45 (75) | 339 (91) | **P < 0.0001***** |
| Mucinous Adenocarcinoma | 30 (7) | 13 (22)*** | 17 (5) | |
| Signet Ring Cell Adenocarcinoma | 16 (4) | 2 (3) | 14 (4) | |
| **TUMOUR GRADE** | **371** | **45** | **326** | |
| Low Grade | 346 (93) | 40 (89) | 306 (94) | P = 0.1022 |
| Medium/High Grade | 25 (7) | 5 (11) | 20 (6) | |
| **AJCC TNM STAGING** | **287** | **45** | **242** | |
| I | 11 (4) | 2 (4) | 9 (4) | |
| II | 101 (35) | 17 (38) | 84 (35) | P = 0.6226 |
| III | 158 (55) | 22 (49) | 136 (56) | |
| IV | 17 (6) | 4 (9) | 13 (5) | |
| **TUMOUR INFILTRATING LYMPHOCYTES (TIL)** | **283** | **47** | **236** | |
| None | 180 (64) | 21 (45) | 159 (67) | **P = 0.0045**** |
| Mild-moderate | 103 (36) | 26 (55)** | 77 (33) | |
| **CHROHN'S LIKE INFLAMMATORY RESPONSE** | **283** | **47** | **236** | |
| None | 206 (73) | 34 (72) | 172 (73) | P = 1.0000 |
| Mild-moderate | 77 (27) | 13 (28) | 64 (27) | |
| **LYMPHATIC INVASION** | **346** | **54** | **292** | |
| Absent | 242 (70) | 34 (63) | 208 (71) | P = 0.2582 |
| Present | 104 (30) | 20 (37) | 84 (29) | |
| **POLYPS** | **313** | **49** | **264** | |
| Absent | 210 (67) | 31 (67) | 179 (68) | P = 0.6197 |
| Present | 103 (33) | 18 (33) | 85 (32) | |
| **POLYP SUBTYPE** | **101** | **17** | **84** | **P = 0.0216* TA vs TVA** |
| Hyperplastic Polyp | 3 (3) | 0 (0) | 3 (4) | |
| Pseudopolyp | 2 (2) | 1 (6) | 1 (1) | |
| Sessile serrated Adenoma (SSA) | 1 (1) | 1 (6) | 0 (0) | |
| Tubular Adenoma (TA) | 53 (52) | 4 (23) | 49 (58)* | |
| Tubulovillous Adenoma (TVA) | 42 (42) | 11 (65)* | 31 (37) | |

**Table 1.** Descriptive analysis of CRC cases diagnosed at CMJAH between (2011–2015) with Microsatellite Instability status, stratified by CRC subtypes: MSI versus MSS.

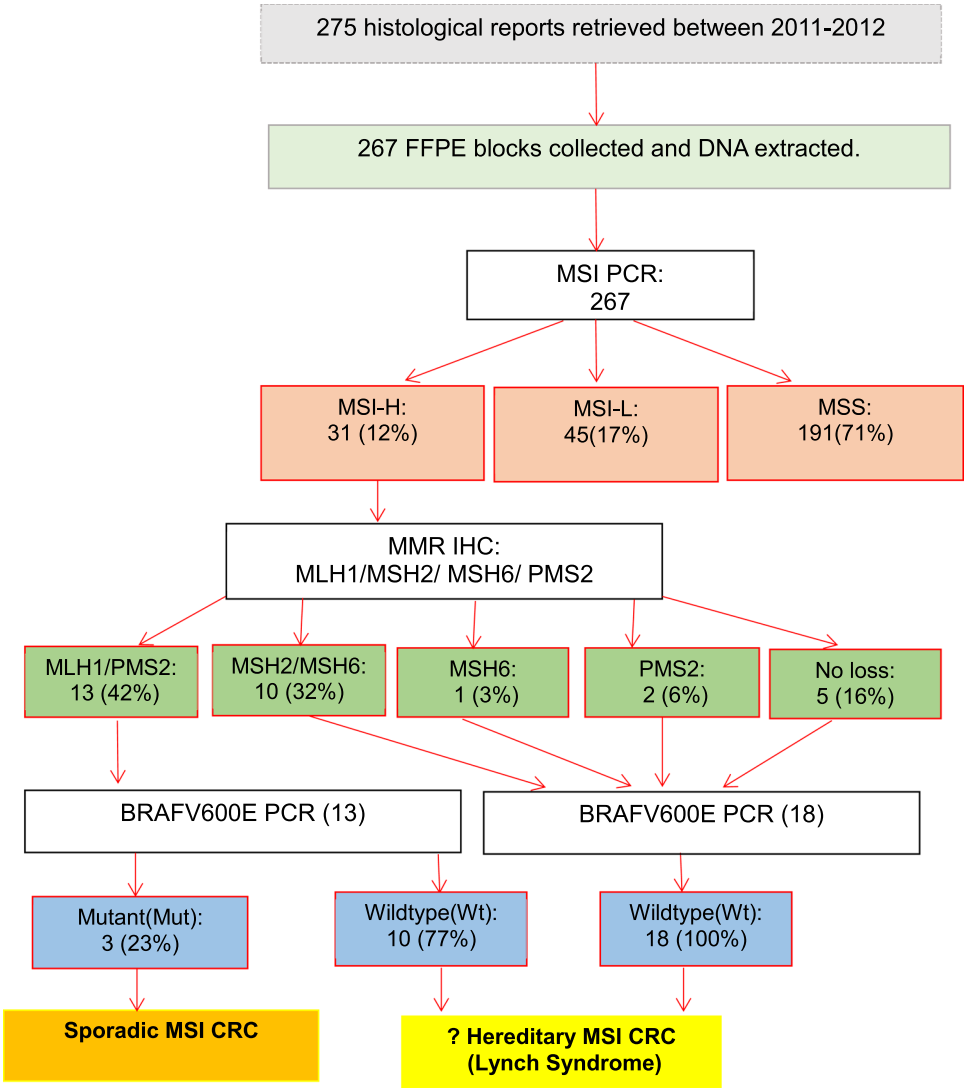



**Figure 1.** Flow diagram illustrating the molecular methodology and frequency results of CRC cases diagnosed between 2011–2012. MSI CRC was found in 12% of CRC cases, with predominant MLH1 loss, however only 23% of deficient MLH1 cases harboured BRAFV600E mutations. These findings suggest a large proportion of SA MSI CRC patients could develop through a hereditary pathway.

adenocarcinomas. Absence of a TIL response was 60% most likely not to have MSI in comparison to those with TILs. TVAs were not included due to the small sample size, reducing the power and value of the analysis.

## Discussion

MSI CRC was detected in 12% of CRC patients diagnosed during the period 2011–2012, however when stratified by ethnic groups, a higher frequency was found in black versus OEG patients although only of borderline significance (15% versus 8% respectively; $P = 0.0831$). Black patients with MSI tumours occurred in younger patients with left- (36%) and right-sided (59%) tumours, associated with dMSH2/6 (45%), and dMLH1/PMS2 (36%) tumours with a lack of BRAF mutations, indicating a possible LS cause[10,33–35]. MSI CRC in Western population groups are shown to dominantly occur in female patients with right-sided tumours, a serrated morphology and BRAFV600E mutations, MLH1 protein loss due to promoter methylation and a high CIMP molecular phenotype[2,20].

In the SA cohort, MSI CRC was associated with younger age, right sided tumours, mucinous adenocarcinomas, TIL response, and TVA precursor lesions. This correlates well with findings in the literature[36,37]. Histopathological features were further stratified by ethnic groups (black versus OEG) to assess if any associations occurred with regard to gender, age, tumour site, MMR protein expression profile and BRAF mutation status. Within the 2011–2012 cohort, patients within the OEG were predominantly older (median age: 62), female (67%), possessed right sided tumours (89%), with deficient MLH1 (56%) expression, of which 60% harboured BRAF mutations. The 2011–2015 cohort however demonstrated an equal ratio of male to female patients, which could signify an under-diagnosis of female patients with MSI within the OEG. All other variables illustrated

| | | MSI CRC stratified by ethnic groups (2011–2015) | | |
|---|---|---|---|---|
| | | No. of cases (%) | | |
| **Demographic Data** | **Number of cases (%)** | **Black** | **OEGs** | **Statistical Significance** |
| **GENDER** | **61** | **39** | **22** | |
| Male | 34 (56) | 23 (59) | 11 (50) | P = 0.5946 |
| Female | 27 (44) | 16 (41) | 11 (50) | |
| **AGE** | **61** | **39** | **22** | |
| Median | 51 | 47*** | 62 | |
| Min-Max | 27–77 | 28–74 | 27–77 | |
| Mean ± SD | 52 ± 13 | 47 ± 11 | 59 ± 14 | **P = 0.0006*** ** |
| P25-P75 | 40–62 | 38–55 | 48–70 | |
| 95% CI | [48–55] | [44–51] | [53–65] | |
| **Categorical Age** | **61** | **39** | **22** | |
| ≤50 years | 29 (48) | 23 (59)* | 6 (27) | **P = 0.0316*** |
| >50 years | 32 (52) | 16 (41) | 16 (73)* | |
| **TUMOUR SITE** | **60** | **39** | **21** | |
| Left | 12 (20) | 10 (26) | 2 (10) | |
| Right | 47 (78) | 28 (72) | 19 (90) | P = 0.1817 Left vs Right |
| Left and Right | 1 (2) | 1 (2) | 0 (0) | |
| **MMR PROTEIN IHC** | **61** | **39** | **22** | |
| dMLH1/PMS2 | 31 (51) | 16 (41) | 15 (68)** | |
| dMSH2/MSH6 | 22 (36) | 19 (49)** | 3 (14) | **P = 0.0098**** MLH1/ PMS2 vs MSH2/MSH6 |
| dMSH6 | 1 (2) | 1 (2.5) | 0 (0) | |
| dPMS2 | 2 (3) | 1 (2.5) | 1 (4) | |
| pMMR | 5 (8) | 2 (5) | 3 (14) | |
| ***BRAF V600E** | **34** | **24** | **10** | |
| Wildtype | 30 (88) | 24 (100)** | 6 (60) | **P = 0.0045**** |
| Mutation | 4 (12) | 0 (0) | 4 (40)** | |

**Table 2.** Descriptive analysis of MSI CRC diagnosed at CMJAH (2011–2015), stratified by ethnic groups: Black versus Other Ethnic Groups (OEGs). Left (Colon distal from splenic flexure); Right (Colon proximal to splenic flexure); Microsatellite Instability (MSI); Mismatch repair (MMR); Deficient (d); Proficient (p). Significance indicated by an asterix. *Majority of the samples screened for BRAFV600E mutations (31/34; 91%) were from the 2011–2012 cohort. The remaining 3 were retrieved from histological reports for the remaining period (2013–2015).

increased statistical associations within the 5-year cohort study. Collectively these features suggest the majority of OEG patients follow the more well established sporadic MSI CRC pathway[19,20]. Within the black patient group, all 2011–2012 data analysed with statistical significant associations revealed increased significance within the 5 year cohort, with the exception of tumour site where 2011–2012 data illustrated significance with an increased incidence of left-sided MSI tumours. This could possibly be due to pathologists generally screening more right than left-sided tumours for MSI CRC.

MSI CRC also showed a high percentage of proficient (p) MMR expression (16%) for the universal 2-year screen, despite high concordance rates (95–96%) for MSI PCR versus IHC reported[38]. This could possibly be indicative of aberrant non-functional proteins expressed or an alternative MMR protein involved in the development of MSI in this group of patients. This would suggest the 5 mononucleotide panel for MSI testing is the more sensitive tool for detecting instability in the South African setting. Perhaps this could be explored further for alternative causes of dMMR in these patients. For example, MSH3 loss of expression has recently been shown to be associated with MSI in African American patients[39,40].

Multiple logistic regression analysis was also conducted to control for the effect of confounding factors, young age (≤50 years), right-sided tumours, mucinous adenocarcinomas and TILS were associated with MSI CRC in SA patients.

Limitations of the study included patient selection bias as patients from the 2013–2015 cohort were only selected if an MSI status was available. In addition, only a few patient samples with available data in certain subgroup analysis, such as TNM stage of disease, TILs, polyp subtype could be analysed, and this is mainly due to biopsy samples also included in the cohort. Only a subset of MSI samples (34/61; 56%) for the period 2011–2015 were included for BRAFV600E analysis, these largely from cases between period 2011–2012. This also represents a limitation of the study. The majority of patients diagnosed at CMJAH are of the black population group, therefore selection bias could play a role in the selection effect due to smaller sample numbers in the comparator OEG patient group. Previous MSI studies in SA however, have largely shown MLH1 mutations as the cause of hereditary and sporadic CRC in OEG patients (specifically the coloured and white population groups in the Western and Northern Cape regions of SA)[23,41,42]. A study carried out by Cronje *et al*. in 2009 at CMJAH in Johannesburg

| Variable | Multivariate | |
|---|---|---|
| | Odds ratio with CI | P-Value |
| **Gender** | | |
| Male | 0.834 (0.394, 1.762) | 0.634 |
| **Age Category** | | |
| >50 years | 0.330 (0.146, 0.745) | **0.008**** |
| **Ethnic Group** | | |
| Black | 0.902 (0.411, 1.983) | 0.799 |
| **Tumour Site** | | |
| Right | 7.661 (3.150, 18.630) | **0.000***** |
| **Subtype** | | |
| Invasive Signet ring cell | 0.971 (0.007, 1.249) | 0.074 |
| Invasive Adenocarcinoma | 0.270 (0.099, 0.731) | **0.010*** |
| **Grade** | | |
| LG | 0.535 (0.142, 2.007) | 0.354 |
| **Tumour Infiltrating Lymphocytes** | | |
| None | 0.392 (0.182, 0.845) | **0.017*** |
| **Lymphatic Invasion** | | |
| Present | 0.829 (0.376, 1.825) | 0.642 |

**Table 3.** Multiple logistic regression factors at risk for MSI CRC. Number of observations were 277. Young Patients under 50 years, with right-sided tumours, displaying mucinous features with TILs were more likely to have MSI. Significance indicated by an asterix.

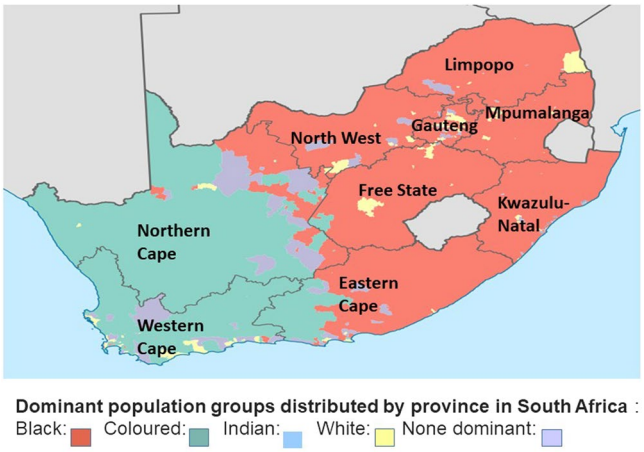



**Figure 2.** Map of South Africa illustrating the distribution of population groups by the different provinces, (adapted from the Wikipedia map of South Africa 2011)[54].

demonstrated a large number of young black patients had increased loss of MMR protein expression, particularly MSH2, in comparison to their white counterpart, also suggesting a heritable cause[21]. Findings from this study confirm dMLH1 and dMSH2 associations within young OEG and black SA patients respectively.

South Africa is comprised of 9 provinces, with an estimated mid-year population of 58.78 million. The Northern Cape comprises 2.2% of the population, Western Cape 11.5%, and Gauteng 25.8%[43,44]. The Western and Northern Cape comprises mainly of the coloured population within SA, and LS have largely been identified in these 2 provinces (See Fig. 2)[23,41,45]. Genetic services are offered to newly diagnosed LS patients and family members in these regions mainly through academic centres based in the Western Cape, and to smaller surrounding towns and rural areas through outreach programmes[23,41,45,46].

The Gauteng province, although the smallest, is the most densely populated and urbanized with nearly 700 people per square kilometre[43,44]. This is largely due to the discovery of gold in Johannesburg during the 1800's,

with a large influx of migrants from Africa and Europe[46]. The black population is the predominant group (77.4%) followed by white (15.6%) in this region[43,44]. LS genetic screening in Gauteng and remaining 6 provinces have been poor to non-existent, even though preliminary findings have suggested a possible hereditary cause, particularly in black patients from the 1990's[21,47–49].

The healthcare financial system in SA is based on private and public funding, with the majority (57%) of the population utilizing public healthcare services, however 80% of healthcare expenditure in SA occurs in the private sector[46,50]. Genetic services such as testing and counselling are mostly offered in academic centres, and subsidies for these services are limited primarily due to the high medical costs associated with the HIV/AIDS epidemic[46]. Pathologists within the National Health Laboratory Services include a possible hereditary cause in histopathological reports of these patients with suspected LS, but genetic services are limited to counselling, particularly within the public sector of the Gauteng province. Colonoscopic surveillance of family members at risk for LS is difficult and expensive in a resource-poor country, however previous LS studies carried out within the Western Cape region of SA have shown the identification of the causative-mutation simplifies and reduces costs of surveillance considerably[22,45,51].

If LS is the more prevalent MSI subtype in black SA patients, identifying the causative mutation together with regular colonoscopy screening is recommended for all affected individuals and families by the 2nd decade of life, to be repeated every 1–2 years[52]. LS patients treated with aspirin have a considerably reduced risk and reduced mortality rates associated with CRC[27,53].

## Conclusion

This study revealed MSI in black SA CRC patients is associated with earlier age onset, higher frequency of MSH2/6 loss and a lack of BRAFV600E mutations, suggesting Lynch or "Lynch-Like Syndrome" (LLS) as the main hereditary driver. Implementation of universal MSI screening is recommended as a routine diagnostic test within the SA setting to increase the detection rate of LS or LLS. This data warrants further molecular investigations in defining the underlying germline or somatic mutations. Appropriate early detection, surveillance and prevention strategies will reduce the risk and improve overall management of the disease in suspected LS patients and affected family members, particularly in the black population group in SA.

**Compliance with ethical standards.** This was a retrospective study on already available biological material and ethics clearance was obtained from the Human Ethics Research Committee, Faculty of Health Sciences: Ethics clearance reference number: M120994.

## Data availability

The dataset generated and analyzed during the current study are available from the corresponding author on reasonable request.

Received: 9 May 2019; Accepted: 17 September 2019;

Published online: 21 October 2019

## References

1. TCGA. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
2. Muller, M., Ibrahim, A. & Arends, M. Molecular pathological classification of colorectal cancer. *Virchows Arch.* **469**, 125–134 (2016).
3. Mojarad, E. N., Kuppen, P. J. K., Aghdaei, H. A. & Zali, M. R. The CpG island methylator phenotype (CIMP) in colorectal cancer. *Gastroenterol. Hepatol. from Bed to Bench* **6**, 120–128 (2013).
4. Muller, A. & Ibrahim, M. A. New molecular classification of colorectal cancer. *Cancer Chemother. Rev.* **11**, 66–72 (2016).
5. Tariq, K., Ghias, K., Tariq, K. & Ghias, K. Colorectal cancer carcinogenesis: a review of mechanisms. *Cancer Biol. Med.* **13**, 120–35 (2016).
6. Wang, L. *et al.* BRAF mutations in colon cancer are not likely attributable to defective DNA mismatch repair. *Cancer Res.* **63**, 5209–5212 (2003).
7. Toyota, M. *et al.* CpG island methylator phenotype in colorectal cancer. *Proc. Natl. Acad. Sci. USA* **96**, 8681–8686 (1999).
8. Kawakami, H., Zaanan, A. & Sinicrope, A. MSI testing and its role in the management of colorectal cancer. *Curr. Treat. Options Oncol.* **16**, 1–12 (2015).
9. Adar, T. *et al.* A tailored approach to BRAF and MLH1 methylation testing in a universal screening program for Lynch syndrome. *Mod. Pathol.* **30**, 440–447 (2017).
10. Weisenberger, D. *et al.* CpG Island Methylator Phenotype in Human Colorectal Cancer Is Tightly Associated with BRAF Mutation and Underlies Sporadic Mismatch Repair Deficiency. *Nat. Genet.* **38**, 787–793 (2006).
11. Lynch, P. M., Lynch, H. T. & Harris, R. E. Hereditary proximal colonic cancer. *Dis. Colon Rectum* **20**, 661–668 (1977).
12. Lynch, H. *et al.* Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin. Genet* **76**, 1–18 (2009).
13. Boland, C. & Goel, A. Microsatellite Instability in Colorectal Cancer. *J. Gastroenterol.* **138**, 2073–2087 (2010).
14. Liccardo, R., De Rosa, M., Izzo, P. & Duraturo, F. Novel implications in molecular diagnosis of lynch syndrome. *Gastroenterol. Res. Pract.* **2017**, 1–12 (2017).
15. Okamoto, K. *et al.* Clinicopathological characteristics of serrated polyps as precursors to colorectal cancer: Current status and management. *J. Gastroenterol. Hepatol.* **32**, 358–367 (2017).
16. Jass, R. *HNPCC and sporadic MSI-H colorectal cancer: a review of the morphological similarities and differences.* **3**, 93–100 (2004).
17. Walsh, M. D. *et al.* Immunohistochemical testing of conventional adenomas for loss of expression of mismatch repair proteins in Lynch syndrome mutation carriers: A case series from the Australasian site of the colon cancer family registry. *Mod. Pathol.* **25**, 722–730 (2012).
18. Boland, C. R. *et al.* 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 Res.* **58**, 5248–5257 (1998).
19. Sinicrope, F. A. & Yang, Z. J. Prognostic and predictive impact of DNA mismatch repair in the management of colorectal cancer. Futur. *Oncol.* **7**, 467–474 (2011).

20. Inamura, K. Colorectal cancers: An update on their molecular pathology. *Cancers (Basel).* **10**, 1–18 (2018).
21. Cronjé, L., Paterson, A. C. & Becker, P. J. Colorectal cancer in South Africa: A heritable cause suspected in many young black patients. *South African Med. J.* **99**, 103–106 (2009).
22. Stupart, D. A., Goldberg, P. A., Algar, U. & Ramesar, R. Surveillance colonoscopy improves survival in a cohort of subjects with a single mismatch repair gene mutation. *Color. Dis.* **11**, 126–130 (2009).
23. Vergouwe, F. *et al*. Mismatch repair deficiency in colorectal cancer patients in a low-incidence area. *South African J. Surg.* **51**, 16–21 (2013).
24. Katsidzira, L. *et al*. The shifting epidemiology of colorectal cancer in sub-Saharan Africa. Lancet. *Gastroenterol. Hepatol.* **2**, 377–383 (2017).
25. Boland, P. M., Yurgelun, M. B. & Boland, C. R. Recent progress in Lynch syndrome and other familial colorectal cancer syndromes. *CA. Cancer J. Clin.* **68**, 217–231 (2018).
26. Sargent, D. J. *et al*. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J. Clin. Oncol.* **28**, 3219–3226 (2010).
27. Burn, J. *et al*. Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: An analysis from the CAPP2 randomised controlled trial. *Lancet* **378**, 2081–2087 (2011).
28. Hendry, S. *et al*. Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors: A Practical Review for Pathologists and Proposal for a Standardized Method from the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in Melanoma, Gastrointestinal Tract Carcinom. *Adv. Anat. Pathol.* **24**, 311–335 (2017).
29. Haghighi, M. M. *et al*. Frequent MSI mononucleotide markers for diagnosis of hereditary nonpolyposis colorectal cancer. *Asian Pac. J. Cancer Prev.* **11**, 1033–1035 (2010).
30. Pyatt, R. *et al*. Implications for Microsatellite Instability Testing. *Cancer Res.* **155**, 349–353 (1999).
31. Suraweera, N. *et al*. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* **123**, 1804–1811 (2002).
32. Buhard, O. *et al*. Multipopulation Analysis of Polymorphisms in Five Mononucleotide Repeats Used to Determine the Microsatellite Instability Status of Human Tumors. *J. Clin. Oncol.* **24**, 241–251 (2005).
33. Umar, A. *et al*. *Revised Bethesda Guidelines for Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndrome) and Microsatellite Instability.* **96**, 261–268 (2004).
34. Lynch, H. & de la Chapelle, A. *Hereditary Colorectal Cancer.* **348**, 919–932 (2003).
35. Lagerstedt Robinson, K. *et al*. Lynch syndrome (hereditary nonpolyposis colorectal cancer) diagnostics. *J. Natl. Cancer Inst.* **99**, 291–299 (2007).
36. Alexander, J. *et al*. Histopathological identification of colon cancer with microsatellite instability. *Am. J. Pathol.* **158**, 527–535 (2001).
37. Kim, H., Jen, J., Vogelstein, B. & Hamilton, S. R. Clinical and pathological characteristics of sporadic colorectal carcinomas with DNA replication errors in microsatellite sequences. *Am J Pathol* **145**, 148–156 (1994).
38. Cicek, M. S. *et al*. Quality assessment and correlation of microsatellite instability and immunohistochemical markers among population- And clinic-based colorectal tumors: Results from the colon cancer family registry. *J. Mol. Diagnostics* **13**, 271–281 (2011).
39. Ashktorab, H. *et al*. Driver genes exome sequencing reveals distinct variants in African Americans with colorectal neoplasia. *Oncotarget* **10**, 2607–2624 (2019).
40. Venderbosch, S. *et al*. EMAST is associated with a poor prognosis in microsatellite instable metastatic colorectal cancer. *PLoS One* **10**, 1–13 (2015).
41. Wentik, M. Q., Rakers, M., Stupart, D. A., Ramesar, R. & Goldberg, P. A. Incidence and histological features of colorectal cancer in the Northern Cape province, South Africa. *South African J. Surg.* **48**, 109–113 (2010).
42. Hitchins, M. *et al*. Identification of new cases of early-onset colorectal cancer with an MLH1 epimutation in an ethnically diverse South African cohort. *Clin. Genet.* **80**, 428–434 (2011).
43. Census. Statistics South Africa. *Census Stat. South Africa* 78 (2011), P0301.4 (2011).
44. Statistics South Africa. Mid-year population estimates. Mid-year Popul. Estim. 2019 (Report). 29 July 2019 Retrieved 29 July 2019 (2019).
45. Bruwer, Z., Futter, M. & Ramesar, R. A mobile colonoscopic unit for lynch syndrome: Trends in surveillance uptake and patient experiences of screening in a developing country. *J. Genet. Couns.* **22**, 125–137 (2013).
46. Kromberg, J. G., Sizer, E. B. & Christianson, A. L. Genetic services and testing in South Africa. *J. Community Genet.* **4**, 413–423 (2013).
47. Degiannis, E., Sliwa, K., Levy, R., Hale, M. J. & Saadia, R. Clinicopathological trends in colorectal carcinoma in a Black South African population. *Trop. Gastroenterol.* **16**, 55–61 (1995).
48. Skalsky, J. A. Colorectal cancer in young black patients. *Trop. Doct.* **19**, 185 (1989).
49. Bebington, B. An evaluation of epidemiological differences in colorectal cancer patients over and under the age of 50 presenting to academic centers in Johannesburg. *S Afr J Surg* **55**, 54 (2017).
50. Econex. Estimating the Financial Cost of the NHI Plan. *Heal. (San Fr.* **9**, 24–30 (2010).
51. Goldberg, P. A. *et al*. In a Resource-Poor Country, Mutation Identification has the Potential to Reduce the Cost of Family Management for Hereditary Nonpolyposis Colorectal Cancer. *Dis. Colon Rectum* **41**, 1250–1255 (1998).
52. Lynch, H. T., Lynch, J. F. & Attard, T. A. Diagnosis and management of hereditary colorectal cancer syndromes: Lynch syndrome as a model. *Cmaj* **181**, 273–280 (2009).
53. Chan, A. T. *et al*. Aspirin in the chemoprevention of colorectal neoplasia: An overview. *Cancer Prev. Res.* **5**, 164–178 (2012).
54. Wikipedia Map of South Africa by dominant population groups. Available at: https://en.wikipedia.org/wiki/Demographics_of_South_Africa#/media/File:South_Africa_2011_dominant_population_group_map.svg, (Accessed: 30th July 2019) (2011).


## Acknowledgements

Griffin Trust Fund (GTCP016). National Health Laboratory Services Trust Fund (94347). Wits University FRC Fund (001.251.8462101.5121105.000000.0000000000.4550). Dr. Innocent Maposa (Biostatistician) for statistical analysis contribution. Alain Lobandji and Nzeba Madi for assisting with the immunohistochemistry methodology.

## Author contributions

M.M., Y.P. and C.P. were involved in the design and research methodology of the study. M.M. performed the analysis and wrote the manuscript. R.M. and P.M. contributed to data analysis. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available this paper at https://doi.org/10.1038/s41598-019-51316-4.

**Correspondence** and requests for materials should be addressed to M.M.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019

# Chapter 4:

Left-sided Colorectal Cancer distinct in Indigenous African patients compared to Other Ethnic Groups in South Africa.

McCabe *et al. BMC Cancer* (2022) 22:1089
https://doi.org/10.1186/s12885-022-10185-3

BMC Cancer

RESEARCH

Open Access

# Left-sided colorectal cancer distinct in indigenous African patients compared to other ethnic groups in South Africa

Michelle McCabe[1,2*], Clement Penny[3], Pumza Magangane[1], Sheefa Mirza[3] and Yvonne Perner[1]

## Abstract

**Introduction:** A large proportion of indigenous African (IA) colorectal cancer (CRC) patients in South Africa are young (< 50 years), with no unique histopathological or molecular characteristics. Anatomical site as well as microsatellite instability (MSI) status have shown to be associated with different clinicopathological and molecular features. This study aimed to ascertain key histopathological features in microsatellite stable (MSS) and low-frequency MSI (MSI-L) patients, to provide insight into the mechanism of the disease.

**Methods:** A retrospective cohort (2011–2015) of MSS/MSI-L CRC patient samples diagnosed at Charlotte Maxeke Johannesburg Academic Hospital was analyzed. Samples were categorized by site [right colon cancer (RCC) versus left (LCC)], ethnicity [IA versus other ethnic groups (OEG)] and MSI status (MSI-L vs MSS). T-test, Fischer's exact and Chi-square tests were conducted.

**Results:** IA patients with LCC demonstrated an increased prevalence in males, sigmoid colon, signet-ring-cell morphology, MSI-L with BAT25/26 marker instability and advanced disease association.

**Conclusion:** This study revealed distinct histopathological features for LCC, and suggests BAT25 and BAT26 as negative prognostic markers in African CRC patients. Larger confirmatory studies are recommended.

**Keywords:** Colorectal cancer, African, Microsatellite stable, Low level microsatellite instability, BAT25, BAT26

## Introduction

Right-sided colon cancer (RCC) and left-sided colon cancer (LCC) show distinct mechanisms of development and are associated with different clinicopathological features [1–3]. During embryological development, the RC develops from the midgut and the LC from the hindgut, supporting the view that RCC and LCC develop through different developmental/embryological pathways genetic mechanisms [1, 4]. The incidence of RCC (~30%) is lower than LCC (~70%). RCC presents with larger tumours, a higher rate of tumour node metastases (TNM stage), mucinous features, and comprises of an overall poorer survival than LCC [1, 2, 4–6]. Older female patients are at higher risk of developing RCC compared to younger male patients associated with increased risk of developing LCC [1]. The literature shows population groups with a lower risk of developing CRC moving to high-risk areas acquire the risks associated with the new area, and this could be linked to dietary, environmental, cultural and genetic factors [1, 7]. This speaks to a possible role for the increasing incidence of CRC in indigenous African (IA) patients moving from rural to urbanized areas [8–10].

*Correspondence: michelle.mccabe@nhls.ac.za

[1] Division of Anatomical Pathology, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, National Health Laboratory Services, Johannesburg 2193, South Africa
Full list of author information is available at the end of the article

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

There are 3 main pathways involved in the development of CRC: 1) Microsatellite instability (MSI) caused by a defective mismatch repair (MMR) system, most often (~70–95%) caused by an alteration of MLH1,2) Chromosome instability (CIN) pathway which develops due to gross chromosomal changes and 3) the CpG island methylator phenotype (CIMP) pathway, arising through methylation of CpG islands in promoter sequences, leading to inactivation of tumour suppressor genes throughout the genome [11, 12]. MSI and CIMP tumours mostly occur in the right colon, whereas CIN CRC is associated with LCC [11, 13, 14]. Four main consensus molecular subtypes (CMS) were established in 2015. These are differentiated by unique molecular features: CMS1 (14%, MSI pathway, immune activation); CMS2 (37%, Canonical WNT/MYC pathway, epithelial signature); CMS3 (13%, epithelial and metabolic dysregulation), and CMS4 (23%; Mesenchymal TGF- β pathway; stromal invasions and angiogenesis) [11, 15].

To date, fewer research outputs on low frequency microsatellite instability (MSI-L) CRC have been published compared to MSI-H and MSS CRC. MSI-L is usually grouped with MSS CRC, as literature states all CRCs display some level of MSI [16, 17]. Some researchers interpret MSI-L tumours as precursors of MSI-H CRC, whereas others believe it to be a completely separate entity [18, 19]. MSI-L tumours have illustrated different clinicopathological features and have been considered to be a worse prognostic group in a few CRC studies [20–23]. Jass et al. reported that MSI-L LCC showed distinctive clinicopathological features, with a male predilection, a moderately differentiated histopathological grade, KRAS mutations, CIMP-Low status and DNA aneuploidy. In contrast, MSI-L RCC was found to occur more frequently in females, being associated with a serrated adenoma precursor lesion, mucinous adenocarcinoma histological subtype and poorly differentiated grade. BRAF mutations, CIMP-High status and diploid DNA content, features associated with a worse prognosis, were also associated with the MSI-L RCC group [24].

A disproportionate number of indigenous African (IA) patients display a younger age of onset (< 50 years of age) with no distinct histopathological features to assist with early diagnosis and management [25–28]. Previous work by McCabe et al. described MSI-H CRC in detail according to ethnicity groups and found an increased association of MSH2/MSH6 MMR protein expression loss in right sided CRC in young IA patients [29]. This study aims to characterize proficient MMR (MSS/MSI-L) CRC, by ethnicity (IA versus OEG) and anatomical site (LCC versus RCC), to potentially identify a unique subtype associated with young IA CRC patients.

## Methodology

### Patient demographics and tumour histopathological characterization

This retrospective study was conducted on a 5-year cohort (2011–2015) of 428CRC patient samples with known anatomical site and MMR status, diagnosed at the Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) branch of the National Health Laboratory Service (NHLS). Informed consent was waived by the Human Research Ethics Committee (HREC) (Medical) of the University of the Witwatersrand for this study, as research was conducted under the institutional blanket ethics clearance (M10744) obtained from the HREC (Medical), which allows for research to be carried out on all archived pathology specimens without informed consent from study participants. Additional project-specific ethical clearance was also obtained from the HREC (Medical) (M120994), and all tests were performed according to the relevant guidelines and regulations. A total of 59 (14%) patient samples had a deficient MMR (dMMR) profile (MSI-H) and the remaining 369 (86%) a proficient MMR (pMMR) status. All pMMR CRC cases were categorized into 4 groups (see Table 1). Tumour site the main group: RCC (tumour primary site proximal to splenic flexure) versus LCC (tumour primary site distal from splenic flexure), further sub-grouped by ethnicity: Indigenous African (IA) versus Other Ethnic Groups (OEG) [Caucasian, Mixed Ancestry, Asian]. Demographic and histopathological information were analyzed within these categories i.e., gender, age, tumour subtype, grade, site, TNM stage (American Joint Committee on Cancer [AJCC] TNM stage), presence of tumour infiltrating lymphocytes (TIL) using the Klintrup-Mäkinen scoring assessment, Crohn's-like inflammatory reaction (CLR), polyp subtype, venous, perineural and lymphatic invasion, all of which were obtained from histology reports. Immunophenotypic profiling of TIL was not conducted. No family histories were available from these reports.

### MSS versus MSI-L molecular subtyping

Samples were screened for proficient MMR status through MMR immunohistochemistry (IHC) and/or MSI polymerase chain reaction (PCR). IHC images and MSI electropherograms are illustrated in Figs. 1 and 2 respectively. The MMR IHC panel included antibodies targeting MutL Homolog 1(MLH1), MutS Homolog 2 and 6 (MSH2/ MSH6) and Post Meiotic Segregation Homolog-2 (PMS2) protein expression. Only samples with a MMR proficient profile detected via IHC and a MSS or MSI-L profile determined via PCR were included in this cohort [29]. MSI PCR included the 5-mononucleotide PCR panel (NR27, NR21, NR24, BAT25 and

**Table 1** Descriptive analysis of MSS CRC cases diagnosed at CMJAH between (2011–2015). Categorized by site: LCC *vs* RCC and Ethnicity: Indigenous African (IA) *vs* Other Ethnic Group (OEG)

| | | Total number cases (%) | Left-sided Colon Cancer (LCC) | | Statistical analysis: | Right-sided Colon Cancer (RCC) | | Statistical analysis: |
|---|---|---|---|---|---|---|---|---|
| | | | IA | OEG | | IA | OEG | |
| **Frequency** | | **369** | **154 (42)** | **108(29)** | | **63(17)** | **44(12)** | |
| **Demographical data** | | | | | | | | |
| **GENDER** | | **369** | **154** | **108** | ***P*=0.0111*** | **63** | **44** | ***P*=0.6932** |
| **Male** | | **209(57)** | **98(64)*** | **51(47)** | | **34(54)** | **26(59)** | |
| **Female** | | **160(43)** | **56(36)** | **57(53)** | | **29(46)** | **18(41)** | |
| **AGE** | | **365** | **152** | **107** | ***P*<0.0001***** | **62** | **44** | ***P*=0.0103*** |
| **Min–Max** | | **15–92** | **20–90** | **28–92** | | **15–79** | **25–86** | |
| **Mean ± SD** | | **57 ± 14** | **53 ± 15** | **62 ± 13** | | **54 ± 13** | **60 ± 11** | |
| **Median** | | **59** | **54***** | **62***** | | **55*** | **61*** | |
| **P25-P75 (Interquartile Range)** | | **47–67** | **41–65** | **55–72** | | **47–64** | **54–68** | |
| **95% CI** | | **[55-58]** | **[51-56]** | **[60-65]** | | **[51-57]** | **[57-65]** | |
| **Categorical Age** | | **365** | **152** | **107** | ***P*<0.0001***** | **62** | **44** | ***P*<0.0001***** |
| **≤ 50 years** | | **107 (29)** | **62(41)***** | **15(14)** | | **24(39)***** | **6(14)** | |
| **> 50 years** | | **258(71)** | **90(59)** | **92(86)** | | **38(61)** | **38(86)** | |
| ***Histological characteristics*** | | | | | | | | |
| **TUMOUR SUBTYPE** | | **360** | **152** | **104** | ***P*=0.0257*** | **61** | **43** | ***P*=0.2608** |
| **Invasive Adenocarcinoma** | | **329(91)** | **138(91)** | **100(96)** | | **54(89)** | **35(81)** | |
| **Mucinous Adenocarcinoma** | | **17(5)** | **4(3)** | **4(4)** | | **3(5)** | **6(14)** | |
| **Signet Ring Cell Adenocarcinoma** | | **14(4)** | **10(6)** | **0(0)** | | **4(6)** | **2(5)** | |
| **TUMOUR SITE:** | | **359** | **151** | **105** | ***P*=0.0221*** | **61** | **42** | ***P*=0.2431** |
| **Left** | **Right** | | | | | | | |
| **Splenic Flexure** | **Hepatic Flexure** | **14(4)** | **5(3)** | **0(0)** | | **6(10)** | **3(7)** | |
| **Descending colon** | **Ascending Colon** | **58(16)** | **21(14)** | **11(10)** | | **11(18)** | **15(36)** | |
| **Sigmoid** | **Transverse Colon** | **83(23)** | **46(31)*** | **21(20)** | | **10(16)** | **6(14)** | |
| **Rectum** | **Caecum** | **204(57)** | **79(52)** | **73(70)** | | **34(56)** | **18(43)** | |
| **TUMOUR GRADE** | | **348** | **148** | **95** | ***P*=0.0930** | **61** | **44** | ***P*=1.0000** |
| **Low Grade (LG)** | | **316(91)** | **132(89)** | **91(96)** | | **54(89)** | **39(89)** | |
| **High Grade (HG)** | | **32(9)** | **16(11)** | **4(4)** | | **7(11)** | **5(11)** | |
| **AJCC TNM STAGING** | | **240** | **85** | **61** | ***P*=0.0922** | **57** | **37** | ***P*=0.4995** |
| **I-II** | | **92(38)** | **31(38)** | **31(51)** | | **20(36)** | **10(27)** | |
| **III-IV** | | **148(62)** | **54(62)** | **30(49)** | | **37(64)** | **27(73)** | |
| **TUMOUR INFILTRATING LYMPHOCYTES (TIL)** | | **230** | **81** | **60** | ***P*=0.8512** | **54** | **35** | ***P*=0.6591** |
| **None** | | **156(68)** | **58(72)** | **44(73)** | | **34(63)** | **20(57)** | |
| **Mild-moderate** | | **74(32)** | **23(28)** | **16(27)** | | **20(37)** | **15(43)** | |
| **CROHN'S LIKE INFLAMMATORY RESPONSE** | | **230** | **81** | **60** | ***P*=0.4406** | **54** | **35** | ***P*=0.3392** |
| **None** | | **168(73)** | **62(77)** | **42(70)** | | **41(76)** | **23(66)** | |
| **Mild-moderate** | | **62(27)** | **19(23)** | **18(30)** | | **13(24)** | **12(34)** | |
| **LYMPHATIC INVASION** | | **285** | **115** | **76** | ***P*=0.2921** | **56** | **38** | ***P*=0.2896** |
| **Absent** | | **203(71)** | **85(74)** | **62(82)** | | **36(64)** | **20(53)** | |
| **Present** | | **82(29)** | **30(26)** | **14(18)** | | **20(36)** | **18(47)** | |
| **VENOUS INVASION** | | **232** | **82** | **59** | ***P*=0.3937** | **55** | **36** | ***P*=1.0000** |
| **Absent** | | **181(78)** | **68(83)** | **45(76)** | | **41(75)** | **27(75)** | |
| **Present** | | **51(22)** | **14(17)** | **14(24)** | | **14(25)** | **9(25)** | |
| **PERINEURAL INVASION** | | **241** | **91** | **60** | ***P*=0.1110** | **54** | **36** | ***P*=1.0000** |
| **Absent** | | **187(78)** | **67(74)** | **51(85)** | | **41(76)** | **28(78)** | |
| **Present** | | **54(22)** | **24(26)** | **9(15)** | | **13(24)** | **8(22)** | |
| **POLYPS** | | **80** | **31** | **19** | ***P*=0.5516** | **16** | **14** | ***P*=0.7131** |
| **Tubular Adenoma (TA)** | | **49(58)** | **21(68)** | **11(58)** | | **10(63)** | **7(50)** | |
| **Tubulovillous Adenoma (TVA)** | | **31(37)** | **10(32)** | **8(42)** | | **6(37)** | **7(50)** | |

**Table 1** (continued)

| | Total number cases (%) | Left-sided Colon Cancer (LCC) | | Statistical analysis: | Right-sided Colon Cancer (RCC) | | Statistical analysis: |
|---|---|---|---|---|---|---|---|
| | | IA | OEG | | IA | OEG | |
| **MSS/MSI-L PCR CONFIRMED CASES** | **233** | **94** | **85** | ***P*=0.0157*** | **28** | **26** | ***P*=0.3582** |
| **MSS** | **189(81)** | **72(77)** | **77(91)** | | **19(70)** | **21(81)** | |
| **MSI-L** | **44(19)** | **22(23)** | **8(9)** | | **9(30)** | **5(19)** | |
| **MSI PCR MARKERS (MSI-L)** | **42** | **22** | **8** | ***P*=0.0138*** | **7** | **5** | ***P*=0.2222** |
| **1 single unstable marker** | | | | **BAT25/26** | | | **BAT25/26** |
| **BAT25** | **20(56)** | **11(50)*** | **5(62.5)** | **Vs** | **3((43)** | **1(20)** | ***Vs*** |
| **BAT26** | **15(22)** | **11(50)*** | **0(0)** | **NR21/24/27** | **3(43)** | **1(20)** | **NR21/24/27** |
| **NR21** | **4(12)** | **0(4)** | **2(25)** | | **0(0)** | **2(40)** | |
| **NR24** | **3(10)** | **0(0)** | **1(12.5)** | | **1(14)** | **1(20)** | |
| **NR27** | **0(0)** | **0(0)** | **0(0)** | | **0(0)** | **0(0)** | |

Levels of statistical significance indicated by an asterix *. *P*-values <0.05, is indicated with (*), *P*-values <0.01, (**), and *p*-values < than 0.001 (***)

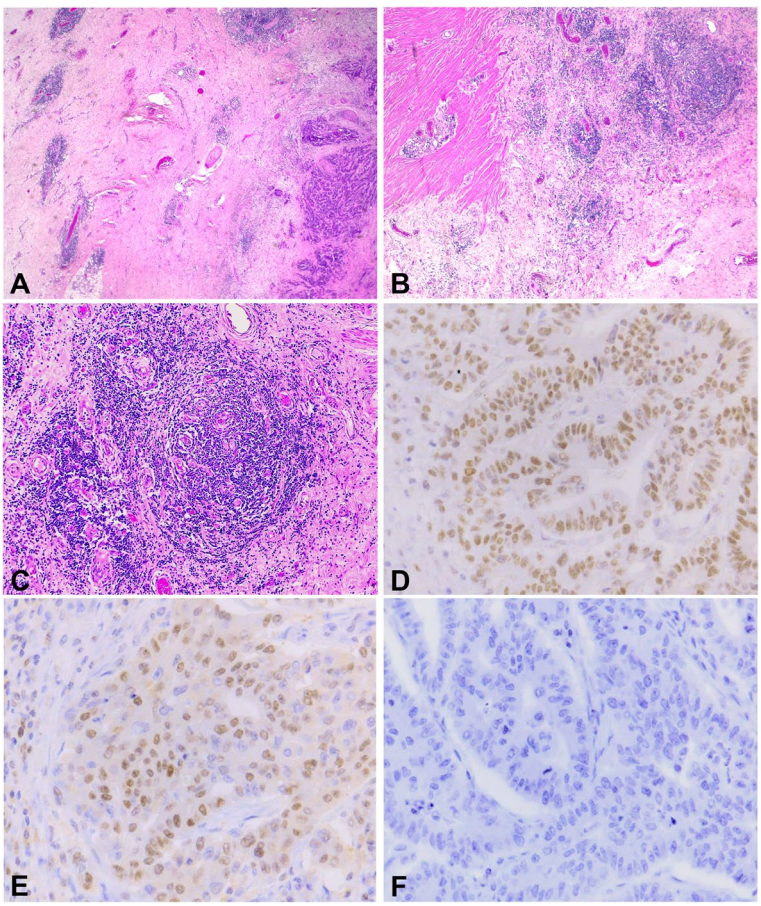



**Fig. 1** Haematoxylin and Eosin (**A-C**) images and Immunohistochemistry (**D-F**) images of colorectal (CRC) tissue sections. **A** Low power (40X) of serosa showing CRC and mild Crohn's-like response (CLR). **B** Low (100X) power of serosa with moderate CLR. **C** Similar CLR at higher magnification (200X). **D** MSH2 retained nuclear expression (400X). **E** MLH6 retained nuclear expression (400X). **F** MLH1 loss of nuclear expression, showing MSIphenotype (400X)

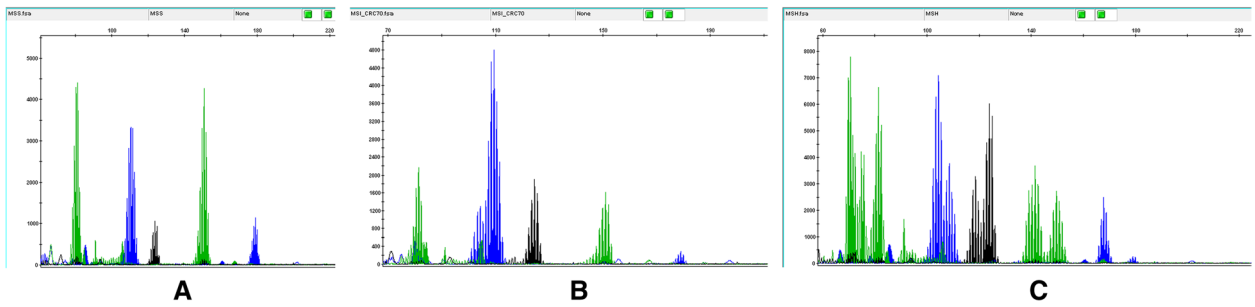



**Fig. 2** Electropherograms illustrating instability patterns using the 5-mononucleotide MSI PCR panel. **A** MSS: 5 markers stable **B** MSI-Low: 1 marker unstable **C** MSI-H: 5 markers unstable

BAT26); and a MSS or MSI-L result was ascribed when an allelic size varied in none or only one of the 5 markers respectively [29, 30]. The literature indicates that people of African descent exhibit normal variation within loci BAT25 and BAT26 [31–34]. A multipopulation study by Buhard et al. 2006, revealed approximately 10% show normal variation in one of five markers and 2% in 2 markers [33]. In silico analysis of BAT 25 and BAT 26 PCR primer sets were performed against the global 1000 Genomes dataset (www.internationalgenome.org/data) and local AWI-GEN dataset (https://www.wits.ac.za/research/sbimb/research/awi-gen). Data analysis from the 1000 Genomes dataset revealed 2 single nucleotide polymorphisms (SNPs) in each primer set and occurred at a rare frequency of 0.02%. AWI-GEN data analysis showed one SNP for BAT26 primers at position Chr2: 47,641,434, and occurred at a frequency of 0.5%. No SNPs were detected in BAT25 primer set. The 5-mononucleotide panel remained the assay of choice due to the additional 3 quasimonomorphic mononucleotide repeat markers found in Caucasian and African germ-line DNA, ensuring the panel is extremely sensitive in detecting somatic alterations in MSI-H tumours and distinguishing between MSS/MSI-L tumours. As recommended by the authors Suraweera et al. 2002, in tumour samples with instability in BAT25 and/or 26 markers with a proficient MMR profile via IHC, matched normal samples were assessed to establish the true instability status. In cases where these markers matched instability in normal colon tissue, the status was regarded as normal or germline variation and reported as MSS. In biopsies this was not possible due to the limited size and mixture of normal and neoplastic tissue, increasing the chance of contamination when assessing normal tissue.

### Statistical analysis

All data was collected in an excel sheet and statistical analysis was performed using Stata Intercooled 7.0 (Stata, College Station, TX, USA) and Graphpad Prism version 9.0 (Graphpad Software, La Jolla, CA, USA). Unpaired t-tests were used to assess differences between groups on continuous normally distributed variables while Fischer's exact and Chi-square tests were used to assess associations between categorical variables; and a result with a $P$ value less than 0.05 was considered statistically significant. Additionally, analysis of MSI status (MSI-L versus MSS and MSI-H CRC) stratified by tumour site (LCC versus RCC) was conducted to determine if any association occurred with demographic variables (gender, age and ethnic groups) as well as TNM stage, CLR and TIL (Table 2). R/Rstudio was used to perform multiple comparisons on MSI status stratified by tumour site with other categorical variables eg gender to ascertain pairwise associations and the False Discovery Rate (FDR) adjusted $p$-value (q-value) was reported and considered statistically significant if less than 0.05.

## Results

### Patient demographics

LCC revealed a higher frequency in males in IA than in OEG patients (64% vs 47%, respectively; $P=0.0111$) (Table 1). The IA patients were younger in comparison to the OEG patients (median age: 54 vs 62 years, respectively; $P<0.0001$).

### Pathological characterization

Signet ring cell carcinomas (SRCC) were more frequently found in IA (14/213; 7%) versus OEG patients (2/147; 1%) ($P=0.0221$). When further stratified by site, SRCC was only associated with LCC in IA patients, as compared to OEG patients ($P=0.0257$).

### MSI-L, MSS and MSI-H molecular subtypes

Ethnicity was linked to mononucleotide instability markers, with BAT25 and BAT26 markers being more

**Table 2** Multiple comparison analysis of CRC cases diagnosed at CMJAH (2011–2015). Categorized by MSI status: MSI-L *vs* MSS *vs* MSI-H CRC and site: LCC *vs* RCC

| | Number of cases (%) | MSI-L CRC | | MSS CRC | | MSI-H CRC | | Statistical analysis: Multiple comparison analysis: *(Associations between 2 groups italicized)* |
|---|---|---|---|---|---|---|---|---|
| | CRC 2011–2015 | A: LCC | B: RCC | C: LCC | D: RCC | E: LCC | F: RCC | |
| **Frequency/ Prevalence** | **303** | **30(10)** | **12(4)** | **160(53)** | **42(14)** | **12(4)** | **47(15)** | |
| **GENDER** | 303 | 30 | 12 | 160 | 42 | 12 | 47 | $X^2 = 2.6877$, df = 5 |
| Male | 160(53) | 16(53) | 6(50) | 88(55) | 24(57) | 6(50) | 20(43) | $P = 0.748$ |
| Female | 143(47) | 14(47) | 6(50) | 72(45) | 18(43) | 6(50) | 27(57) | *No association of sex* |
| **AGE** | 303 | 29 | 12 | 160 | 42 | 12 | 47 | FDR (q value = 0.8238) Q = 0.05 |
| Min–Max | 15–92 | 24–91 | 37–84 | 20–92 | 15–80 | 33–67 | 27–77 | |
| Mean ± SD | 57 ± 14 | 56 ± 17 | 59 ± 13 | 58 ± 15 | 58 ± 13 | 50 ± 12 | 53 ± 14 | *E vs C: P = 0.0415** |
| Median | 59 | 56 | 60 | 59 | 58 | 52 | 51 | *F vs C: P = 0.0274** |
| P25-P75 (IQR) | 48–69 | 40–72 | 47–67 | 48–69 | 53–67 | 38–61 | 43–62 | |
| 95% CI | [56-59] | [49-62] | [50–67] | [55-60] | [53-62] | [42-57] | [49-57] | |
| **ETHNIC GROUPS** | 303 | 30 | 12 | 160 | 42 | 12 | 47 | $X^2 = 9.2824$, df = 5, $P = 0.09832$ |
| Indigenous African | 171(56) | 22(76) | 7(58) | 83(47) | 21(50) | 10(83) | 28(60) | *A vs C: P = 0.0442** |
| Other Ethnic groups | 132 (44) | 8(24) | 5(42) | 77(53) | 21(50) | 2(17) | 19(40) | *E vs C: P = 0.0393** |
| **AJCC TNM STAGING** | 178 | 9 | 8 | 76 | 37 | 7 | 41 | $X^2 = 7.6681$, df = 5, $P = 0.1755$ |
| I-II | 77(43) | 1(11) | 4(50) | 38(50) | 12(32) | 4(57) | 18(44) | |
| III-IV | 101(57) | 8(89) | 4(50) | 38(50) | 25(68) | 3(43) | 23(56) | *A vs C: P = 0.0348** |
| **CROHN'S LIKE INFLAMMATORY RESPONSE** | 172 | 9 | 8 | 73 | 35 | 7 | 40 | $X^2 = 11.185$, df = 5, $P = 0.04782$* |
| Mild-moderate | 49(28) | 6(67) | 1(12) | 19(26) | 10(29) | 4(57) | 9(23) | *A vs C: P = 0.0205** |
| Absent | 123(72) | 3(33) | 7(88) | 54(74) | 25(71) | 3(43) | 31(77) | |
| **TUMOR INFILTRATING LYMPHOCYTES** | 178 | 15 | 8 | 73 | 35 | 7 | 40 | $X^2 = 12.898$, df = 5, $P = 0.02435$* *D vs C: P = 0.0497** |
| Mild-moderate | 68(38) | 5(33) | 2(25) | 19(26) | 16(46) | 5(71) | 21(52) | *E vs C: P = 0.0231** |
| Absent | 110(62) | 10(67) | 6(75) | 54(74) | 19(54) | 2(29) | 19(48) | *F vs C: P = 0.0072*** |

Levels of statistical significance indicated by an asterix *. *P*-values <0.05 indicated with (*), and *P*-values <0.01 with (**)

frequently unstable in IA patients (28/29; 97%) within the MSI-L subgroup. NR21, NR24 and N27 instability was commonly demonstrated in OEG patients (7/13; 54%) ($P = 0.0053$). An increased rate of MSI-L vs MSS CRC (21% vs 9%; $P = 0.0442$) and MSI-H vs MSS CRC (11% vs 2.5%; $P = 0.0393$) was found in the left colon, particularly in IA compared to OEG patients (Table 2). In addition, the MSI-L subtype was associated with more advanced disease stage (III-IV) (8/9; 89%) when compared to MSS CRC (38/76; 50%) ($P = 0.0348$). CLR was associated with MSI-L LCC (6/9; 67%), when compared to subtypes MSI-L RCC (1/8; 12%) ($P = 0.0498$), and MSS LCC (19/73; 26%) ($P = 0.0205$). TIL was associated with subtypes MSI-H LCC ($P = 0.0231$), MSI-H RCC ($P = 0.0072$) and MSS-RCC ($P = 0.0497$).

## Discussion

CRC has been shown to have unique clinicopathological features associated with tumour site and different molecular subtypes (CMS 1–4) [11]. CRC molecular subtypes and age of onset have also been described to vary considerably among geographically distinct ethnic groups. Within this cohort, (40%) of IA patients was shown to be younger (< 50 years) compared to OEGs with pMMR CRC. Patients of OEGs in this cohort displayed RCC with poor prognostic factors compared to LCC. Increased frequencies of HG tumours (11% vs 4%) advanced staged tumours (73% vs 49%), perineural invasion (22% vs 15%), mucinous and signet ring morphology (19% vs 4%) were seen in RCC. While the median age of onset was similar for LCC vs RCC (61 vs 62), more males however were diagnosed with RCC (59%) than LCC (47%).

Within the IA patient group, both left and right colon cancers showed similar frequencies for poor prognostic factors. Higher frequencies for HG tumours (11%), advanced tumour stage (62–64%), perineural invasion (24–26%), SRCC (6%), younger age onset (median age 54–55), with more males presenting with LCC compared to RCC (64% vs 54%).

When comparing ethnic groups and right versus left-sided CRC, significant differences were observed for IA

patients with LCC. The IA population group showed a propensity to occur in males, within the sigmoid colon, to present with a SRCC histological pattern, and an MSI-L status. Notably, SRCC is recognized as a rare histological subtype (1%) of CRC and is associated with young adults in other geographical locations [35]. Previous studies have shown SRCC to have a RCC dominance. However, more recently, SRCC has been reported to have an even site distribution within the colon, with a slight male predominance [36]. Moreover, the SRCC histological subtype is known to have an adverse prognostic significance independent of tumour stage and molecular subtype [37, 38]. Poor tumour grade and advanced TNM staging are usually associated with worse survival outcomes [39, 40]. In this study, these features were found to have borderline significance in IA patients with LCC compared to OEGs, with slight increases in frequencies of HG (11% vs 4%, respectively; $P=0.0930$) and advanced disease stage (62% vs 49%, respectively; $P=0.0922$).

When assessing PCR confirmed cases only to accurately categorize MSI-L versus MSS and MSI-H CRC in right versus left CRC, significant associations were seen for MSI-L LCC with advanced disease stage and the IA ethnic group (Table 2). This data was perceived to be similar to the findings of Devaraj 2010, linking elevated microsatellite alterations at selected tetranucleotide repeats (EMAST) with advanced disease stage, rectal cancers and peri-tumoral infiltration in patients of African American (AA) descent [41, 42]. Even though MSI-L status was not available for the data in the study of Devaraj et al. 2010, previously published data closely linked MSI-L tumours to EMAST and has been associated with a poorer prognosis [20, 22, 43]. The most frequent unstable markers in tumours from IA patients in our cohort were BAT25 (15/29; 52%) and BAT26 (13/29; 45%). MSI-L LCC tumours showed a tendency to be of a more advanced disease stage (AJCC TNM stage: III-IV) compared to MSS LCC (89% vs 50%, respectively; $P=0.0348$).

Previous studies have shown these markers to be polymorphic within the African population, linked to the theory that older population groups show increased genetic variation [31–33]. Within this cohort, 17 CRC patients had shown instability in one or both markers (BAT25 and/or 26) in tumour and matched normal tissue, with a proficient MMR protein expression profile. These samples were assumed to be due to increased genetic polymorphisms. What was interesting and noteworthy to mention, was all these patients were exclusively of IA descent, with no differences observed in allele deletion sizes (5-15 bp deletion) between germline and somatic instability, and the majority were LCC (15/17; 88%) with advanced disease stage at diagnosis (7/10; 70%). Based on these findings, limiting as it is in size, BAT25/26 instability (whether of polymorphic/germline or somatic variation) was associated with advanced disease stage in proficient MMR LCC patients of IA descent. Even though survival data was not available for this study, literature has shown poor clinical prognosis and overall survival associated with MSI-L CRC, particularly for advanced disease stage CRC [21–23, 44]. The somatic MSI-L LCC group was associated with CLR when compared to MSI-L RCC (67% vs 12%; $P=0.0498$) and MSS LCC (67% vs 26%; $P=0.0205$). Polymorphic/germline MSI-L tumours within 2 markers (11/17; 65%) however displayed no CLR. Literature has indicated tumours with CLR to have a better prognosis compared to stage-matched tumours without [45–47]. This raises a plausible argument that germline MSI-L tumours could have a worse prognosis compared to somatic MSI-L tumours, due to the lack of the host's immune response to the cancer.

BAT26 marker (26(A) repeats) is located in intron 5 of the MSH2 gene on chromosome 2p21. This marker is situated immediately downstream of exon 5, which is susceptible to large intragenic deletions and accounts for nearly a third of dMSH2 mutations [48–50]. Studies by Pastrello et al. 2006 and Jaskowski et al. 2007 indicated that instability in Bat26 was associated with overall instability of dMSH2 tumours. Confirmatory IHC to determine dMMR protein expression is therefore important, however exceptions of cases with mutations in intronic nucleotides close to splice sites could result in expressed non-functional proteins, such as (*MSH2* c.913G>A p.Ala305Thr) which has been reported with proficient MMR activity and a MSI-L genotype [50]. This variant however had no aberrant splicing and normal subcellular localization and interaction with MSH6 was shown [50, 51].

BAT25 (25(T) repeats) is situated within intron 16 of the c-kit proto-oncogene on chromosome 4q12. cKit (CD117) the receptor for Stem cell factor (SCF) involved in haemopoiesis has more recently shown to be involved in lymphopoiesis. The CD117 receptor has shown to be expressed on mature $CD8^+$ T cells following initial activation, suppressing differentiation and increasing its response to apoptosis. CD117 expressed $CD8^+$ T cells could therefore play a role in CD117-blockade, an important mechanism in tumour immune evasion. BAT25 instability in the CD117 gene could potentially play a role in immune evasion in MSI-L CRC, and additional studies are required to determine CLR and its association with MSI-L LCC.

A study by Carethers et al. 2014 showed an increased incidence of MSI-H CRC in AA patients, however had poorer prognosis and higher mortality rates compared to their Caucasian counterparts [52]. This was thought to be

due to AA patients showing a lower infiltration of $CD8^+$ T cells compared to Caucasian patients, suggesting an altered immune function in AA patients. It has been well established that the increased tumour-infiltrating $CD4^+$ and $CD8^+$ T cells in patients with MSI-H CRC (due to the increased mutator phenotype of the tumour stimulating the host immune response) significantly improved patient outcomes when treated with immunotherapy when compared to MSS CRC with decreased immune response [53, 54]. The study by Carethers et al. 2014 illustrate despite having the same disease subtype and stage, ethnicity can be a negative prognostic factor in CRC disease.

A frequency of 17% MSI-L CRC in SA CRC patients has been reported in our previous study [29]. Further evaluation in this study has demonstrated MSI-L LCC to occur predominantly in IA patients, and associated with advanced disease stage, with considerable number of germline/polymorphic MSI-L LCC also presenting at an advance stage compared to MSI-L RCC and MSS LCC.

Based on these findings, universal MSI PCR screening is recommended as a first-line screening method for all newly diagnosed CRC patients, to not only identify MSI-H CRC, but also increase the detection rate of MSI-L CRC. Local ethnic polymorphisms however have to be taken into account when implementing diagnostic marker panels in certain geographical settings. If instability is required in 30% of markers used in a panel for a diagnosis of MSI, it is important to confirm markers included are non-polymorphic in the local population.

CRC is a heterogeneous disease and more studies are required to unravel the complexity associated with it by investigating different ethnic groups in the context of site. This study has shown that ethnicity and tumour site play an important role in the prognostication of tumours and should be taken into consideration for effective treatment planning, especially in geographical regions with diverse population groups such as South Africa. Limitations of this study include selection bias, as only samples with an MSI status were included. This resulted in smaller sample sizes for the analysis of certain features such as polyps, TILs, TNM staging, MSI-L and BAT25/26 instability status. The lack of universal screening for MSI within the study institution, as well as the inclusion of biopsy samples in addition to resections, have contributed to providing limited information. Confirmation of MSI-L status in biopsies was not possible, increasing the likelihood of a small percentage of false positive MSI-L samples with normal variation.

A study by Ozaki et al. found that MSI-L colon tissue occurred in a few but not all intestinal crypts, and both in malignant and normal tissues. The presence of MSI-L in non-neoplastic mucosa could indicate a primary step in tumorigenesis and could potentially be used as an early diagnostic and prognostic marker in CRC [55].

Additional AA patient studies have illustrated increased frequencies of MSI-L/EMAST markers in rectal cancers most likely due to somatic inactivation of an alternative MMR gene (MSH3) [21, 44, 56, 57]. Dysfunctional MSH3 has shown to lead to MSI, appearing to be inflammation-related within the tumour microenvironment [58]. Regular intake of anti-inflammatory drugs such as aspirin and non-steroidal anti-inflammatory drugs (NSAID) has been reported to prevent the development of colorectal adenomas, tumour growth and progression, as well recurrence and metastasis after curative surgery, prolonging colorectal cancer patient survival [59–61]. Anti-inflammatories could therefore possibly have a positive effect not only in MSI-H CRC, but also in this subgroup of MSI-L LCC patients. In addition, due to the presence of CLR, MSI-L LCC could potentially be an eligible subgroup for immunotherapeutic strategies in metastatic disease and further studies are recommended.

## Conclusion

This SA CRC study indicated that in considering categorization of CRC according to anatomical site, microsatellite instability status and ethnicity, unique clinicopathological features were identified. In particular, IA CRC patients with LCC are more likely to be male, have an MSI-L subtype, show BAT25/BAT26 marker instability and have advanced disease stage. This study suggests that BAT25 and BAT26 instability are negative prognostic markers in African CRC patients, and larger confirmatory studies are recommended. Further exploratory studies of MSH3, EMAST, KRAS and immune cell infiltration in the tumour microenvironment are indicated in SA CRC patients. This will assist in establishing molecular profiles to accurately improve diagnostic, prognostic and personalized predictive markers for the effective management of early onset CRC.

#### Acknowledgements

Dr Jean-Tristan Brandenburg for primer insilico data analysis, Dr. Innocent Maposa (Biostatistician) for statistical analysis contribution, and Mr Eric Liebenberg for immunohistochemistry image capturing.

#### Authors' contributions

MM, YP and CP were involved in the design and research methodology of the study. MM conducted the tests, performed the analysis and wrote the manuscript. All authors reviewed the manuscript. The author(s) read and approved the final manuscript.

#### Funding

Griffin Trust Fund (GTCP016).
National Health Laboratory Services Trust Fund (94,347).
Wits University FRC Fund (4550).
National Research Foundation (001.403).

**Availability of data and materials**
The dataset generated and analysed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Ethics approval and consent to participate**
This was a retrospective study on already available biological material and ethical clearance was obtained from the Human Ethics Research Committee (Medical), University of the Witwatersrand Ethics clearance reference number: M120994. The need for informed consent was waived by the HREC (Medical), University of the Witwatersrand for this study, as research was conducted under the institutional blanket ethics clearance (M10744) obtained from the HREC (Medical), which allows for research to be carried out on all archived pathology specimens without informed consent from study participants. All experimental protocols were approved by the Postgraduate Research and Protocol committee within the Faculty of Health Sciences, University of the Witwatersrand, and all experiments were performed according to the relevant guidelines and regulations.

**Consent for publication**
Not applicable.

**Competing interests**
The authors have no conflicts of interest to declare.

**Author details**
[1]Division of Anatomical Pathology, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, National Health Laboratory Services, Johannesburg 2193, South Africa. [2]Division of Human Genetics, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, National Health Laboratory Services, Braamfontein, Johannesburg 2000, South Africa. [3]Department of Internal Medicine, Faculty of Health Sciences, University of the Witwatersrand, Parktown, Johannesburg 2193, South Africa.

Received: 25 August 2021 Accepted: 14 October 2022
Published online: 24 October 2022

**References**

1. Iacopetta B. Are there two sides to colorectal cancer? Int J Cancer. 2002;101:403–8.
2. Mik M, Berut M, Dziki L, Trzcinski R, Dziki A. Right-and left-sided colon cancer-clinical and pathological differences of the disease entity in one organ. Arch Med Sci. 2017;13(1):157–62.
3. Lim DR, Kuk JK, Kim T, Shin EJ. Comparison of oncological outcomes of right-sided colon cancer versus left-sided colon cancer after curative resection. Med (United States). 2017;96:1–7.
4. Benedix F, Kube R, Meyer F, Schmidt U, Gastinger I, Lippert H. Comparison of 17,641 patients with right- and left-sided colon cancer: differences in epidemiology, perioperative course, histology, and survival. Dis Colon Rectum. 2010;53:57–64.
5. Holch JW, Ricard I, Stintzing S, Modest DP, Heinemann V. The relevance of primary tumour location in patients with metastatic colorectal cancer: A meta-analysis of first-line clinical trials. Eur J Cancer. 2017;70:87–98.
6. Baran B, Mert Ozupek N, Yerli Tetik N, Acar E, Bekcioglu O, Baskin Y. Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. Gastroenterol Res. 2018;11:264–73.
7. Medhanie GA, Fedewa SA, Adissu H, DeSantis CE, Siegel RL, Jemal A. Cancer incidence profile in sub-Saharan African-born blacks in the United States: Similarities and differences with US-born non-Hispanic blacks. Cancer. 2017;123:3116–24.
8. Walker ARP, Segal I. Colorectal cancer in an African city population in transition. Eur J Cancer Prev. 2002;11:187–91.
9. Katsidzira L, Gangaidzo I, Thomson S, Rusakaniko S, Matenga J, Ramesar R. The shifting epidemiology of colorectal cancer in sub-Saharan Africa. Lancet Gastroenterol Hepatol. 2017;2:377–83.
10. Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA Cancer J Clin. 2009;59:366–78.
11. Guinney J, Dienstmann R, Wang X, De Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015;21:1350–6.
12. Kim SY, Kim T Il. Serrated neoplasia pathway as an alternative route of colorectal cancer carcinogenesis. Intest Res. 2018;16:358–65.
13. Stintzing S, Tejpar S, Gibbs P, Thiebach L, Lenz HJ. Understanding the role of primary tumour localisation in colorectal cancer treatment and outcomes. Eur J Cancer. 2017;84:69–80.
14. Samowitz WS, Albertsen H, Herrick J, Levin TR, Sweeney C, Murtaugh MA, et al. Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. Gastroenterology. 2005;129:837–45.
15. Muller M, Ibrahim A, Arends M. Molecular pathological classification of colorectal cancer. Virchows Arch. 2016;469:125–34.
16. Ogino S, Nosho K, Kirkner GJ, Kawasaki T, Meyerhardt JA, Loda M, et al. CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. Gut. 2009;58:90–6.
17. Hile SE, Shabashev S, Eckert KA. Tumor-specific microsatellite instability: Do distinct mechanisms underlie the MSI-L and EMAST phenotypes? Mutat Res - Fundam Mol Mech Mutagen. 2013;743–744:67–77.
18. Mori Y, Selaru FM, Sato F, Yin J, Simms LA, Xu Y, et al. The impact of microsatellite instability on the molecular phenotype of colorectal tumors. Cancer Res. 2003;63:4577–82.
19. Pawlik TM, Raut CP, Rodriguez-bigas MA. Colorectal carcinogenesis : MSI-H versus MSI-L. 2004;20:199–206.
20. Torshizi Esfahani A, Seyedna SY, Nazemalhosseini Mojarad E, Majd A, Asadzadeh Aghdaei H. MSI-L/EMAST is a predictive biomarker for metastasis in colorectal cancer patients. J Cell Physiol. 2018;234:13128–36.
21. Lee SY, Kim DW, Lee HS, Ihn MH, Oh HK, Min BS, et al. Low-level microsatellite instability as a potential prognostic factor in sporadic colorectal cancer. Med (United States). 2015;94:1–8.
22. Mojarad EN, Kashfi SMH, Mirtalebi H, Taleghani MY, Azimzadeh P, Savabkar S, et al. Low Level of Microsatellite Instability Correlates with Poor Clinical Prognosis in Stage II Colorectal Cancer Patients. J Oncol. 2016;2016:1–9.
23. Wright CM, Dent OF, Newland RC, Barker M, Chapuis PH, Bokey EL, et al. Low level microsatellite instability may be associated with reduced cancer specific survival in sporadic stage C colorectal carcinoma. Gut. 2005;54:103–8.
24. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology. 2007;50:113–30.
25. Cronjé L, Paterson AC, Becker PJ. Colorectal cancer in South Africa: A heritable cause suspected in many young black patients. South African Med J. 2009;99:103–6.
26. Prodehl L, Bebington B, Fabian J, Singh E, Ruff P. COLORECTAL CANCER IN A SOUTH AFRICA URBAN SETTING - A PRELIMINARY ANALYSIS. S Afr J Surg. 2017;55:58.
27. Bebington B, Singh E, Fabian J, Jan Kruger C, Prodehl L, Surridge D, et al. Design and methodology of a study on colorectal cancer in Johannesburg, South Africa. JGH Open. 2018;2:139–43.
28. McCabe M, Perner Y, Magobo R, Mirza S, Penny C. Descriptive epidemiological study of South African colorectal cancer patients at a Johannesburg Hospital Academic institution. JGH Open. 2020;4:360–7.
29. McCabe M, Perner Y, Magobo R, Magangane P, Mirza S, Penny C. Microsatellite instability assessment in Black South African colorectal cancer patients reveal an increased incidence of suspected lynch syndrome. Sci Rep. 2019;9:1–10.
30. Haghighi MM, Javadi GR, Parivar K, Milanizadeh S, Zali N, Fatemi SR, et al. Frequent MSI mononucleotide markers for diagnosis of hereditary non-polyposis colorectal cancer. Asian Pac J Cancer Prev. 2010;11:1033–5.
31. Pyatt R, Chadwick RB, Johnson CK, Adebamowo C, Chapelle A De, Prior TW. Implications for microsatellite instability testing. Cancer Res. 1999;155:349–53.
32. Suraweera N, Duval A, Reperant M, Vaury C, Furlan D, Leroy K, et al. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. Gastroenterology. 2002;123:1804–11.
33. Buhard O, Cattaneo F, Wong YF, Yim SF, Friedman E, Flejou J-F, et al. Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. J Clin Oncol. 2005;24:241–51.

34. Brennetot C, Buhard O, Jourdan F, Flejou JF, Duval A, Hamelin R. Mononucleotide repeats BAT-26 and BAT-25 accurately detect MSI-H tumors and predict tumor content: Implications for population screening. Int J Cancer. 2005;113:446–50.
35. Farraj FA, Sabbagh H, Aridi T, Fakhruddin N, Farhat F. Signet Ring Cell Carcinoma of the Colon in Young Adults: A Case Report and Literature Review. Case Rep Oncol Med. 2019;2019:1–8.
36. Wei Q, Wang X, Gao J, Li J, Li J, Qi C, et al. Clinicopathologic and molecular features of colorectal adenocarcinoma with signet-ring cell component. PLoS One. 2016;11:1–12.
37. Hamilton S, Bosman F, Boffetta P, Ilyas M, Morreau H, Nakamura S, et al. Tumours of the colon and rectum: Chapter 8. In: Bosman FT, Carneiro F, Hruban RH, D TN, editors. WHO Classification of Tumours of the Digestive System 2010. 4th ed. IARC: Lyon 2010; 2010. p. 137–8.
38. Hartman DJ, Nikiforova MN, Chang DT, Chu E, Bahary N, Brand RE, et al. Signet Ring Cell Colorectal Carcinoma. Am J Surg Pathol. 2013;37:969–77.
39. Schneider NI, Langner C. Prognostic stratification of colorectal cancer patients: Current perspectives. Cancer Manag Res. 2014;6:291–300.
40. Fleming M, Ravula S, Tatishchev SF, Wang HL. Colorectal carcinoma: Pathologic aspects. J Gastrointest Oncol. 2012;3:153–73.
41. Devaraj B, Lee A, Cabrera BL, Miyai K, Luo L, Ramamoorthy S, et al. Relationship of EMAST and Microsatellite Instability Among Patients with Rectal Cancer. J Gastrointest Surg. 2010;14:1521–8.
42. Lee SY, Miyai K, Han HS, Hwang DY, Seong MK, Chung H, et al. Microsatellite instability, EMAST, and morphology associations with T cell infiltration in colorectal neoplasia. Dig Dis Sci. 2012;57:72–8.
43. Garcia M, Choi C, Kim HR, Daoud Y, Toiyama Y, Takahashi M, et al. Association between recurrent metastasis from stage II and III primary colorectal tumors and moderate microsatellite instability. Gastroenterology. 2012;143:48–50.
44. Kohonen-Corish MRJ, Daniel JJ, Chan C, Lin BPC, Kwun SY, Dent OF, et al. Low microsatellite instability is associated with poor prognosis in stage C colon cancer. J Clin Oncol. 2005;23:2318–24.
45. Buckowitz A, Knaebel HP, Benner A, Bläker H, Gebert J, Kienle P, et al. Microsatellite instability in colorectal cancer is associated with local lymphocyte infiltration and low frequency of distant metastases. Br J Cancer. 2005;92:1746–53.
46. Rozek LS, Schmit SL, Greenson JK, Tomsho LP, Rennert HS, Rennert G, et al. Tumor-Infiltrating lymphocytes, Crohn's-like lymphoid reaction, and survival from colorectal cancer. J Natl Cancer Inst. 2016;108:1–8.
47. Posch F, Silina K, Leibl S, Mündlein A, Moch H, Siebenhüner A, et al. Maturation of tertiary lymphoid structures and recurrence of stage II and III colorectal cancer. Oncoimmunology. 2018;7–13.
48. Pastrello C, Baglioni S, Tibiletti MG, Papi L, Fornasarig M, Morabito A, et al. Stability of BAT26 in tumours of hereditary nonpolyposis colorectal cancer patients with MSH2 intragenic deletion. Eur J Hum Genet. 2006;14:63–8.
49. Jaskowski L, Young J, Jackson L, Arnold S, Barker MA, Walsh MD, et al. Stability of BAT26 in Lynch syndrome colorectal tumours [1]. Eur J Hum Genet. 2007;15:139–41.
50. Arnold S, Buchanan DD, Barker M, Jaskowski L, Michael D, Birney G, et al. Classifying MLH1 and MSH2 variants using bioinformatic prediction, splicing assays, segregation and tumor characteristics Sven. Hum Mutat. 2009;30:757–70.
51. Thompson BA, Martins A, Spurdle AB. A review of mismatch repair gene transcripts: Issues for interpretation of mRNA splicing assays. Clin Genet. 2015;87:100–8.
52. Carethers JM, Murali B, Yang B, Doctolero RT, Tajima A, Basa R, et al. Influence of race on microsatellite instability and CD8+ T cell infiltration in colon cancer. PLoS One. 2014;9:1–7
53. Wang J, Liu J, Tian F, Zhan Y, Kong D. Cyclin-dependent kinase 9 expression and its association with CD8+ T cell infiltration in microsatellite-stable colorectal cancer. Oncol Lett. 2019;18:6046–56.
54. Frumento G, Zuo J, Verma K, Croft W, Ramagiri P, Chen FE, et al. CD117 (c-Kit) is expressed during CD8+ T cell priming and stratifies sensitivity to apoptosis according to strength of TCR engagement. Front Immunol. 2019;10:468.
55. Ozaki K, Nagasaka T, Notohara K, Kambara T, Takeda M, Sasamoto H, et al. Heterogeneous microsatellite instability observed within epithelium of ulcerative colitis. Int J Cancer. 2006;119:2513–9.
56. Ashktorab H, Daremipouran M, Devaney J, Varma S, Rahi H, Lee E, et al. Identification of novel mutations by exome sequencing in African American colorectal cancer patients. Cancer. 2015;121:34–42.
57. Carethers JM, Koi M, Tseng-Rogenski SS. EMAST is a form of microsatellite instability that is initiated by inflammation and modulates colorectal cancer progression. Genes (Basel). 2015;6:185–205.
58. Koi M, Tseng-Rogenski SS, Carethers JM. Inflammation-associated microsatellite alterations: Mechanisms and significance in the prognosis of patients with colorectal cancer. World J Gastrointest Oncol. 2018;10:1–14.
59. Cole BF, Logan RF, Halabi S, Benamouzig R, Sandler RS, Grainge MJ, et al. Aspirin for the chemoprevention of colorectal adenomas: Meta-analysis of the randomized trials. J Natl Cancer Inst. 2009;101:256–66.
60. Arber N, Eagle CJ, Spicak J, Rácz I, Dite P, Hajer J, et al. Celecoxib for the prevention of colorectal adenomatous polyps. N Engl J Med. 2006;355:885–95.
61. Hamada T, Cao Y, Qian ZR, Masugi Y, Nowak JA, Yang J, et al. Aspirin use and colorectal cancer survival according to tumor CD274 (programmed cell death 1 ligand 1) expression status. J Clin Oncol. 2017;35:1836–44.


## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

**Learn more** biomedcentral.com/submissions

BMC

# Chapter 5: Concluding comments

CRC remains a complex disease with multiple environmental and genetic factors involved in its aetiology. The main contributing factors are environmental, with the adoption of a westernized diet and a less active lifestyle. Disease onset predominantly presents from the 7th decade of life, and has been linked to high income countries globally, as elevated incidence rates occur within these regions compared to much lower incidence reported in the under-developed regions, such as Africa.[1,95] To date the incidence of CRC is reported to be decreasing in the developed countries due to effective surveillance and awareness programs in place, however in lower income or developing countries incidences continue to rise.[1] On a global scale however, young onset CRC incidence is increasing, with no clear explanation to the cause in the shift of epidemiology. [96]

Early CRC studies in Africa reported the disease predominantly occurring in the Caucasian population group, and a rare occurrence among Africans.[12,95,97] NCR reports suggest this to be a major underestimate due to inefficient or non-existent cancer registries within rural areas and less developed parts of the country at the time.[24] Alternatively the introduction of an environmental factor such as lifestyle changes (urbanization) or a major genetic founder effect occurring at certain point in time could have contributed to the rise in incidence over the years within the African population.[98–100]

To date the incidence rates of CRC is on the rise in all population groups, and drastic interventions need to be implemented to alleviate the burden of the disease.[9,27] A major concern is the high number of young African patients diagnosed with the disease.[27,28,101] Minimal studies have been published on the prognosis, survival and burden of disease in this patient group.[29,97] The elevated proportion of young onset cancer has been shown in a few studies over the past 2 decades, however no significant strategies or interventions to

determine the causative and underlying genetic factors to assist in the eradication of the disease has been implemented.[101–103]

This study was undertaken to provide further clarity into the molecular subgroups and clinicopathological characteristics of the disease. Overall these findings revealed African CRC patients display unique molecular pathological features compared to other ethnic groups within SA. Universal microsatellite instability screening demonstrated an increase in the detection of MSI CRC patients compared to only using the Bethesda guidelines when selecting patients for MSI testing. The Bethesda criteria for predicting MSI CRC (based on the MSI phenotype, tumour site (RCC), young age of onset and family cancer history), was more restrictive demonstrating a lower pick up rate when comparing to universal MSI PCR screening of SA CRC patients.[104,105] This was due to a proportion of MSI CRC in our population also detected in the left colon and over the age of 50 years. The IHC technique also showed to be less sensitive in comparison to PCR when detecting MSI CRC, this is mainly due to some cases showing expression of non-functional proteins and being detected by IHC. An additional benefit of universal MSI PCR screening is the ability to differentiate MSI-L CRC from the MSS CRC, a subtype demonstrating unique histopathological and molecular features and perceived to be a completely separate molecular entity. MSI genetic polymorphism in 2 markers (BAT25/26) was also associated with advanced disease stage in our cohort. These preliminary results indicated a possibility of BAT25/26 to be utilized as a negative prognostic marker in African CRC patients. This also raises the question whether this genetic polymorphism could be a possible predisposing genetic factor in African patients. Larger confirmatory as well as functional studies are recommended to determine if this 2 marker instability pattern results in an expressed but non-functional protein in these patients. CLR has been associated with this subtype, and literatue has shown improved stage specific prognosis with an immune host response infiltration, and further research into treatment strategies for tumours with increased CLR need to be explored, as illustrated with MSI-H tumours,TIL and immune checkpoint inhibitor therapy.

Futher characterization of all CRC genes, including those providing predictive and prognostic information, such as KRAS variatns and- anti-EGFR therapy should be explored in SA, which has been long overdue and lacking in the effective treatment of the disease, particularly in the public sector.[57,106]

There is a global need to gain greater understanding into the various subtypes of CRC.[41] This includes unravelling molecular signatures of all subtypes to gain insight into the best appropriate targeted treatment and management strategies, to improve patient survival outcomes, quality of life and the overall burden associated with the disease. Considering all research outputs from this study, the implementation of a routine universal MSI PCR screening tool in diagnostic histopathology laboratories is recommended in SA. Subsequent to confirming MSI status, IHC analysis should be performed on all MSI-H CRC to determine the causative defective protein. Additional BRAFV600E testing can be conducted on all *MLH1* deficient tumours to differentiate between sporadic and hereditary forms of the disease. This study highlighted *MSH2* as the predominant causative gene of young onset MSI CRC in the African population, and recommend additional investigation by next generation sequencing to determine the causative pathological variants in somatic and germline cells of deficient MSH2 CRC patients. This will assist in ascertaining whether the dominant MSI molecular subtype in African patients is due to a genetic predisposition or sporadic cause. As there are multiple possible genetic causes of the disease, including single nucleotide variation (SNV), copy number variation (CNV), epigenetic variation etc., next generation sequencing (NGS) screening of intronic and exonic regions of associated CRC genes is recommended as a first line screening method (Figure 1). Larger deletions, such as those seen in the *EPCAM* gene will not be detected by technologies such as NGS, and therefore CNV by MLPA (multiple ligation probe analysis) or larger deletions by microarray analysis should be implemented as a second line screening tool. If there are no pathological variants identified through this method, additional epigenetic analysis of promoter regions in CRC-related genes need to be pursued.

LS CRC studies within the Western Cape identified a founder mutation (*MLH1*:c.1528C>T) in the mixed ancestry (coloured) population group mainly in the Northern Cape of South Africa.[70,107,108109] Identifying this variant allowed a rapid and cost-effective screening tool to be developed to screen and diagnose patients presenting with symptoms of the disease on a large scale. This allowed for appropriate interventions to be implement in this part of the country within the mixed ancestry population.[108,110–112]

In determining the causative pathological variants in the African population, will allow the same appropriate and effective diagnostic, predictive and management strategies to be implemented.

Policy that all suspected hereditary CRC patients be referred by clinicians to the genetics clinic for follow up should be implemented. Greater effort and priority should given in determining causative factors and underlying germline or somatic pathogenic variants in the aetiology of the disease. Once the common or founder variants have been identified in the causative genes, cost effective genetic screening tools can be implemented in routine diagnostic laboratories to aid in the accurate diagnosis of patients.[113] This information will assist geneticists in the early diagnosis and management of patients, as well the implementation of effective prevention strategies by cascase testing of high risk family members of patients. This information will also guide clinicians in stratifying the best treatment and surveillance strategies in CRC patients.

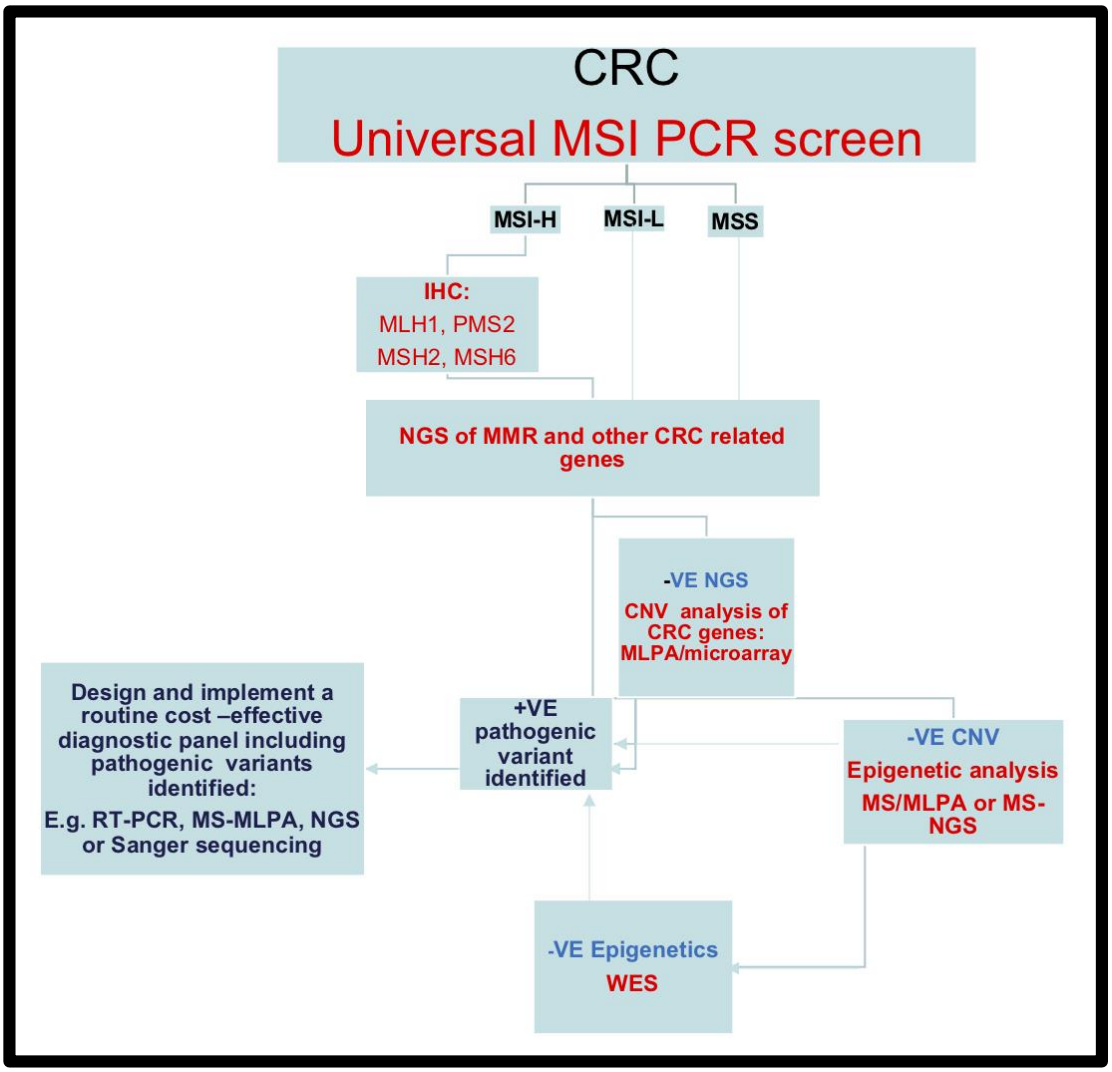



Figure 1: Algorithmic approach recommended for the molecular characterization of CRC in SA

Universal microsatellite instability (MSI) assessment of all CRC tissue samples followed by immunohistochemistry (IHC) of MSI tumours would assist pathologists in effectively determining the molecular subtype of CRC for appropriate treatment and management. The next approach should include next generation sequencing (NGS) utilising targeted CRC-related gene panels for assessment of High-level-MSI (MSI-H), low-level-MSI (MSI-L) and microsatellite stable (MSS) tumours. Both tumour and matched blood specimen types should be screened to establish the underlying genetic cause whether germline or somatic. Samples negative for pathogenic variants by NGS, should be assessed for copy number variation (CNV) through technologies such as multiplex ligation probe analysis (MLPA) and/or microarrays screening larger deletions or duplications. Samples negative for CNV, additional methylation profiling of CRC-related genes should be assessed through methylation specific MLPA or NGS technologies. The final recommended assessment would be whole exome sequencing (WES) of samples with no identifiable genetic cause of disease. Identification of clinically significant variants through this algorithmic approach would assist in the design and implemention of a cost-effective routine diagnostic screening tool for the effective management of CRC patients in SA.

REFERENCES:


1. Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* 71, 209–249 (2021).

2. World Health Organization/ International Agency for Research on Cancer. Cancer Fact Sheets. *Globocan 2020 (IARC)* (2020). Available at: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx. (Accessed: 23 December 2022)

3. Center, M. M., Jemal, A., Smith, R. A. & Ward, E. Worldwide variations in colorectal cancer. *CA Cancer J Clin* 59, 366–78 (2009).

4. Ahmedin Jemal DVM, P. *et al.* Global cancer statistics - Jemal - 2011 - CA A Cancer Journal for Clinicians - Wiley Online Library. *CA: a cancer journal for clinicians* (2011).

5. Ferlay, J. *et al. Cancer incidence and mortality - Major patterns in GLOBOCAN 2012, worldwide and Georgia*. *International Journal of Cancer* 136, E359–E386 (2015).

6. Arnold, M. *et al.* Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 66, 683–691 (2017).

7. Siegel, R., Miller, K. & Jemal, A. Colorectal Cancer Facts & Figures. (2017).

8. Eser, S. *et al.* Incidence patterns of colorectal cancers in four countries of the Middle East Cancer Consortium (Cyprus, Jordan, Israel, and Izmir, Turkey) compared with those in the United States Surveillance, Epidemiology, and End Results Program. *Turkish J. Gastroenterol.* 29, 35–43 (2018).

9. Katsidzira, L. *et al.* The shifting epidemiology of colorectal cancer in sub-Saharan Africa. *Lancet Gastroenterol. Hepatol.* 2, 377–383 (2017).

10. Simon, K. Colorectal cancer development and advances in screening. *Clin. Interv. Aging*

**11**, 967–976 (2016).

11. American Cancer Society. Cancer Facts. *Cancer Facts* 1–9 (2016). doi:10.1097/01.NNR.0000289503.22414.79

12. Graham, A., Adeloye, D., Grant, L., Theodoratou, E. & Campbell, H. Estimating the incidence of colorectal cancer in Sub-Saharan Africa: A systematic analysis. *J. Glob. Health* **2**, (2012).

13. South African National Cancer Registry. Cancer in South Africa (2014). *www.ncr.ac.za* (2014). doi:http://www.cansa.org.za/south-african-cancer-statistics/

14. South African National Cancer Registry. Cancer in South Africa (2015). *www.ncr.ac.za* (2015).

15. South African National Cancer Registry. Cancer in South Africa (2016). *www.ncr.ac.za* (2016). Available at: http://www.cansa.org.za/south-african-cancer-statistics/. (Accessed: 10th July 2018)

16. South African National Cancer Registry. Cancer in South Africa (2011). *www.ncr.ac.za* (2011). Available at: http://www.cansa.org.za/south-african-cancer-statistics/. (Accessed: 25th October 2018)

17. South African National Cancer Registry. Cancer in South Africa (2012). *www.ncr.ac.za* (2012). Available at: http://www.cansa.org.za/south-african-cancer-statistics/. (Accessed: 25th October 2018)

18. South African National Cancer Registry. Cancer in South Africa (2013). *www.ncr.ac.za* (2013). Available at: http://www.cansa.org.za/south-african-cancer-statistics/. (Accessed: 25th October 2018)

19. South African National Cancer Registry. Cancer in South Africa (2006). *www.ncr.ac.za*

(2006). Available at: http://www.cansa.org.za/south-african-cancer-statistics/.

20. South African National Cancer Registry. Cancer in South Africa (2007). *www.ncr.ac.za* 1–38 (2007).

21. South African National Cancer Registry. Cancer in South Africa (2008). *www.ncr.ac.za* (2008). Available at: 2020-12-31.

22. South African National Cancer Registry. Cancer in South Africa (2009). *www.ncr.ac.za* 1–38 (2009).

23. South African National Cancer Registry. Cancer in South Africa (2010). *www.ncr.ac.za* (2010). doi:http://www.cansa.org.za/south-african-cancer-statistics

24. Singh, E. *et al.* South African National Cancer Registry: Effect of withheld data from private health systems on cancer incidence estimates. *S Afr Med J* 105, 107–109 (2016).

25. South African National Cancer Registry. Cancer in South Africa (2017). *www.ncr.ac.za* 1, (2017).

26. South African National Cancer Registry. Cancer in South Africa (2018). *www.ncr.ac.za* 1–38 (2018).

27. South African National Cancer Registry. Cancer in South Africa (2019). *www.ncr.ac.za* 1–38 (2019).

28. Cronjé, L., Paterson, A. C. & Becker, P. J. Colorectal cancer in South Africa: A heritable cause suspected in many young black patients. *South African Med. J.* 99, 103–106 (2009).

29. Prodehl, L. M., Bebington, B., Fabian, J., Singh, E. & Ruff, P. Colorectal Cancer in a South Africa Urban Setting - a Preliminary Analysis. *S Afr J Surg* 55, 58 (2017).

30. Yamagishi, H., Kuroda, H., Imai, Y. & Hiraishi, H. Molecular pathogenesis of sporadic colorectal cancers. *Chin. J. Cancer* **35**, 1–8 (2016).

31. Wells, K. & Wise, P. E. Hereditary Colorectal Cancer Syndromes. *Surg. Clin. North Am.* **97**, 605–625 (2017).

32. Jahani-Sherafat, S., Alebouyeh, M., Moghim, S., Amoli, H. A. & Ghasemian-Safaei, H. Role of gut microbiota in the pathogenesis of colorectal cancer; A review article. *Gastroenterol. Hepatol. from Bed to Bench* **11**, 101–109 (2018).

33. Inamura, K. Colorectal cancers: An update on their molecular pathology. *Cancers (Basel).* **10**, 1–18 (2018).

34. Brim, H. *et al.* A microbiomic analysis in African Americans with colonic lesions reveals streptococcus sp.VT162 as a marker of neoplastic transformation. *Genes (Basel).* **8**, (2017).

35. Ito, M. *et al.* Association of Fusobacterium nucleatum with clinical and molecular features in colorectal serrated pathway. *Int. J. Cancer* **137**, 1258–1268 (2015).

36. Kim, H., Jen, J., Vogelstein, B. & Hamilton, S. R. Clinical and pathological characteristics of sporadic colorectal carcinomas with DNA replication errors in microsatellite sequences. *Am J Pathol* **145**, 148–156 (1994).

37. Lengauer, C., Kinzler, K. & Vogelstein, B. . Genetic instabilities in human cancers. *Nature* **396**, 643–49 (1998).

38. Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).

39. Toyota, M. *et al.* CpG island methylator phenotype in colorectal cancer. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 8681–8686 (1999).

40. Park, B. H. & Vogelstein, B. DNA Repair Pathway Genes DNA Mismatch Repair Gene Defects and HNPCC. 2–4 (2017).

41. Muller, M., Ibrahim, A. & Arends, M. Molecular pathological classification of colorectal cancer. *Virchows Arch.* **469**, 125–134 (2016).

42. Mojarad, E. N., Kuppen, P. J. K., Aghdaei, H. A. & Zali, M. R. The CpG island methylator phenotype (CIMP) in colorectal cancer. *Gastroenterol. Hepatol. from Bed to Bench* **6**, 120–128 (2013).

43. Chen, S. & Shen, X. Long noncoding RNAs: functions and mechanisms in colon cancer. *Mol. Cancer* **19**, 1–13 (2020).

44. Zhu, J. *et al.* MicroRNAs Associated With Colon Cancer: New Potential Prognostic Markers and Targets for Therapy. *Front. Bioeng. Biotechnol.* **8**, 1–10 (2020).

45. David P. Bartel. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function Review. *Cell* **116**, 281–297 (2004).

46. TCGA. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).

47. Muller, A. Ibrahim, M. A. New molecular classification of colorectal cancer. *Cancer Chemother. Rev.* **11**, 66–72 (2016).

48. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**, 1350–1356 (2015).

49. Wikipedia (2011) Map of South Africa by dominant population groups. Available at: https://en.wikipedia.org/wiki/Demographics_of_South_Africa#/media/File:South_Africa_2011_dominant_population_group_map.svg. (Accessed: 30th July 2018)

50. Grady, W. M. Genomic instability and colon cancer. *Cancer Metastasis Rev.* **23**, 11–27

(2004).

51. Pino, M. S. & Chung, D. C. The Chromosomal Instability Pathway in Colon Cancer. *Gastroenterology* **138**, 2059–2072 (2010).

52. Reimers, M. S., Zeestraten, E. C. M., Kuppen, P. J. K., Liefers, G. J. & van de Velde, C. J. H. Biomarkers in precision therapy in colorectal cancer. *Gastroenterol. Rep.* **1**, 166–183 (2013).

53. Fernández-Medarde, A. & Santos, E. Ras in cancer and developmental diseases. *Genes and Cancer* **2**, 344–358 (2011).

54. Zhao, B. *et al.* Mechanisms of resistance to anti-EGFR therapy in colorectal cancer. *Oncotarget* **8**, 3980–4000 (2017).

55. Russo, A. L. *et al.* Mutational analysis and clinical correlation of metastatic colorectal cancer. *Cancer* **120**, 1482–1490 (2014).

56. Jones, R. P. *et al.* Specific mutations in KRAS codon 12 are associated with worse overall survival in patients with advanced and recurrent colorectal cancer. *Br. J. Cancer* **116**, 923–929 (2017).

57. Allegra, C. J. *et al.* Extended RAS gene mutation testing in metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy: American society of clinical oncology provisional clinical opinion update 2015. *J. Clin. Oncol.* **34**, 179–185 (2016).

58. Rock Moon, J., Ju Oh, S., Kyun Lee, C., Gil Chi, S. & Jong Kim, H. TGF-β1 protects colon tumor cells from apoptosis through XAF1 suppression. *Int. J. Oncol.* **54**, 2117–2126 (2019).

59. Itatani, Y., Kawada, K. & Sakai, Y. Transforming growth factor-β signaling pathway in

colorectal cancer and its tumor microenvironment. *Int. J. Mol. Sci.* 20, (2019).

60. Sarshekeh, A. M. *et al.* Association of SMAD4 mutation with patient demographics, tumor characteristics, and clinical outcomes in colorectal cancer. *PLoS One* 12, 1–14 (2017).

61. Haidle, J. L., MacFarland, S. P. & Howe, J. R. Juvenile Polyposis Syndrome. *GeneReviews, Univ. Washington, Seattle* 1–19 (2003).

62. Aubrey, B. J., Kelly, G. L., Janic, A., Herold, M. J. & Strasser, A. How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? *Cell Death Differ.* 25, 104–113 (2018).

63. Schell, M. J. *et al.* A multigene mutation classification of 468 colorectal cancers reveals a prognostic role for APC. *Nat. Commun.* 7, 1–12 (2016).

64. Frebourg, T. *et al.* Germ-line p53 mutations in 15 families with Li-Fraumeni syndrome. *Am. J. Hum. Genet.* 56, 608–615 (1995).

65. Advani, S. M. *et al.* Clinical, Pathological, and Molecular Characteristics of CpG Island Methylator Phenotype in Colorectal Cancer: A Systematic Review and Meta-analysis. *Transl. Oncol.* 11, 1188–1201 (2018).

66. Tahara, T. *et al.* Fusobacterium in colonic flora and molecular features of colorectal carcinoma. *Cancer Res.* 74, 1311–1318 (2014).

67. Boland, C. & Goel, A. Microsatellite Instability in Colorectal Cancer. *J. Gastroenterol.* 138, 2073–2087 (2010).

68. Ma, H. *et al.* Pathology and genetics of hereditary colorectal cancer. *Pathology* 50, (2018).

69. Karanikas, M. & Esebidis, A. Increasing incidence of colon cancer in patients <50

years old: a new entity? *Ann. Transl. Med.* **4**, 164–164 (2016).

70. Hitchins, M. *et al.* Identification of new cases of early-onset colorectal cancer with an MLH1 epimutation in an ethnically diverse South African cohort. *Clin. Genet.* **80**, 428–434 (2011).

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

72. Chan, T. *et al.* Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nat. Genet.* **38**, 1178–83 (2006).

73. Tutlewska, K., Lubinski, J. & Kurzawski, G. Germline deletions in the EPCAM gene as a cause of Lynch syndrome - literature review. *Hered. Cancer Clin. Pract.* **11**, 1–9 (2013).

74. Kang, S. Y. *et al.* Lynch-like syndrome: Characterization and comparison with EPCAM deletion carriers. *Int. J. Cancer* **136**, 1568–1578 (2015).

75. O'Brien, J., Hayder, H., Zayed, Y. & Peng, C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Front. Endocrinol. (Lausanne).* **9**, 1–12 (2018).

76. Fu, G., Brkić, J., Hayder, H. & Peng, C. MicroRNAs in human placental development and pregnancy complications. *Int. J. Mol. Sci.* **14**, 5519–5544 (2013).

77. Xu, W., Lucas, A. S., Wang, Z. & Liu, Y. Identifying microRNA targets in different gene regions. *BMC Bioinformatics* **15(S7)**, 1–11 (2014).

78. Broughton, J., Lovci, M., Huang, J., Yeo, G. & Pasquinelli, A. Mol Cell. *Physiol. Behav.* **64**, 320–333 (2016).

79. Zhang, J. *et al.* Oncogenic role of microRNA-532-5p in human colorectal cancer via targeting of the 5'UTR of RUNX3. *Oncol. Lett.* **15**, 7215–7220 (2018).

80. Dharap, A., Pokrzywa, C., Murali, S., Pandi, G. & Vemuganti, R. MicroRNA miR-324-3p

induces promoter-mediated expression of RelA gene. *PLoS One* **8**, 4–8 (2013).

81. Census. Statistics South Africa. Census (2011). *Stat. South Africa* 78 (2011). doi:P0301.4

82. Hammond, S. An overview of microRNA. *Adv Drug Deliv Rev* **29**, 3–14 (2015).

83. Emmrich, S. *et al.* miR-99a/100~125b tricistrons regulate hematopoietic stem and progenitor cell homeostasis by shifting the balance between TGFβ and Wnt signaling. *Genes Dev.* **28**, 858–874 (2014).

84. Pidíkova, P., Reis, R. & Herichova, I. Mirna clusters with down-regulated expression in human colorectal cancer and their regulation. *Int. J. Mol. Sci.* **21**, 1–22 (2020).

85. Wang, S. *et al.* A plasma microRNA panel for early detection of colorectal cancer. *Int. J. Cancer* **136**, 152–161 (2015).

86. Chen, X. *et al.* Characterization of microRNAs in serum: A novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res.* **18**, 997–1006 (2008).

87. Arroyo, J. D. *et al.* Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5003–5008 (2011).

88. Gallo, A., Tandon, M., Alevizos, I. & Illei, G. G. The majority of microRNAs detectable in serum and saliva is concentrated in exosomes. *PLoS One* **7**, 1–5 (2012).

89. Weber, J. A. *et al.* The microRNA spectrum in 12 body fluids. *Clin. Chem.* **56**, 1733–1741 (2010).

90. Sun, Y. M., Lin, K. Y. & Chen, Y. Q. Diverse functions of miR-125 family in different cell contexts. *J. Hematol. Oncol.* **6**, 1–8 (2013).

91. Liu, H. *et al.* Downregulation of MIR-143 modulates KRAS expression in colorectal

carcinoma cells. *Oncol. Rep.* **42**, 2759–2767 (2019).

92. Li, T. *et al.* MicroRNA-224 sustains Wnt/β-catenin signaling and promotes aggressive phenotype of colorectal cancer. *J. Exp. Clin. Cancer Res.* **35**, 1–11 (2016).

93. Strubberg, A. M. & Madison, B. B. MicroRNAs in the etiology of colorectal cancer: pathways and clinical implications. *DMM Dis. Model. Mech.* **10**, 197–214 (2017).

94. Chirshev, E., Oberg, K. C., Ioffe, Y. J. & Unternaehrer, J. J. Let-7 as biomarker, prognostic indicator, and therapy for precision medicine in cancer. *Clin. Transl. Med.* **8**, (2019).

95. Irabor, D. O. Colorectal carcinoma: Why is there a lower incidence in Nigerians when compared to caucasians? *J. Cancer Epidemiol.* **2011**, (2011).

96. Myers, E. A. *et al.* Colorectal cancer in patients under 50 years of age: A retrospective analysis of two institutions' experience. *World J. Gastroenterol.* **19**, 5651–5657 (2013).

97. Bebington, B. An evaluation of epidemiological differences in colorectal cancer patients over and under the age of 50 presenting to academic centers in Johannesburg. *S Afr J Surg* **55**, 54 (2017).

98. Anderson, W. F., Umar, A. & Brawley, O. W. Colorectal carcinoma in black and white race. *Cancer Metastasis Rev.* **22**, 67–82 (2003).

99. Irabor, D. Emergence of colorectal cancer in West Africa: Accepting the inevitable. *Niger. Med. J.* **58**, 87 (2017).

100. J.C., M. *et al.* Long-term effect of resistant starch on cancer risk in carriers of hereditary colorectal cancer: An analysis from the CAPP2 randomised controlled trial. *Lancet Oncol.* **13**, 1242–1249 (2012).

101. Prodehl, L., Bebington, B., Fabian, J., Singh, E. & Ruff, P. COLORECTAL CANCER IN A

SOUTH AFRICA URBAN SETTING - A PRELIMINARY ANALYSIS. *S Afr J Surg* **55**, 58 (2017).

102. Walker, A. R. P. & Segal, I. Colorectal cancer in an African city population in transition. *Eur. J. Cancer Prev.* **11**, 187–191 (2002).

103. Phipps, A. I. *et al.* Association between molecular subtypes of colorectal cancer and patient survival. *Gastroenterology* **148**, 77-87.e2 (2015).

104. Umar, A. *et al.* NIH Public Access. **96**, 261–268 (2010).

105. McCabe, M. *et al.* Microsatellite Instability assessment in Black South African Colorectal Cancer patients reveal an increased incidence of suspected Lynch syndrome. *Sci. Rep.* **In press**, 1–10 (2019).

106. Sepulveda, A. R. *et al.* Molecular biomarkers for the evaluation of colorectal cancer: Guideline from the American society for clinical pathology, college of American pathologists, association for molecular pathology, and American society of clinical oncology. *J. Clin. Oncol.* **35**, 1453–1486 (2017).

107. Id, L. K. *et al.* Investigation on the hereditary basis of colorectal cancers in an African population with frequent early onset cases. (2019).

108. Burn, J. *et al.* Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: An analysis from the CAPP2 randomised controlled trial. *Lancet* **378**, 2081–2087 (2011).

109. Wentik, M. Q. ., Rakers, M. ., Stupart, D. A. ., Ramesar, R. . & Goldberg, P. A. Incidence and histological features of colorectal cancer in the Northern Cape province, South Africa. *South African J. Surg.* **48**, 109–113 (2010).

110. Stupart, D. A., Goldberg, P. A., Algar, U. & Ramesar, R. Surveillance colonoscopy improves survival in a cohort of subjects with a single mismatch repair gene mutation.

*Color. Dis.* **11**, 126–130 (2009).

111. Vergouwe, F. *et al.* Mismatch repair deficiency in colorectal cancer patients in a low-incidence area. *South African J. Surg.* **51**, 16–21 (2013).

112. Bruwer, Z., Futter, M. & Ramesar, R. A mobile colonoscopic unit for lynch syndrome: Trends in surveillance uptake and patient experiences of screening in a developing country. *J. Genet. Couns.* **22**, 125–137 (2013).

113. Goldberg, P. A. . *et al.* In a Resource-Poor Country, Mutation Identification has the Potential to Reduce the Cost of Family Management for Hereditary Nonpolyposis Colorectal Cancer. *Dis. Colon Rectum* **41**, 1250–1255 (1998).

APPENDIX A:

R14/49 Professor Martin Hale

## HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

## CLEARANCE CERTIFICATE NO. M150884

| | |
|---|---|
| **NAME:** (Principal Investigator) | Professor Martin Hale |
| **DEPARTMENT:** | Division of Anatomical Pathology<br>NHLS |
| **PROJECT TITLE:** | Use of Pathology Specimens Archived in the Department of Anatomical Pathology<br>(Previously M00/08/29/Prof A Paterson)(M10744) |
| **DATE CONSIDERED:** | Adhoc |
| **DECISION:** | Approved unconditionally |
| **CONDITIONS:** | Renewal (2015 - 2020) |
| **SUPERVISOR:** | |
| **APPROVED BY:** | Professor P Cleaton-Jones, Chairperson, HREC (Medical) |
| **DATE OF APPROVAL:** | 09/09/2015 |

**This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.**

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Secretary in Room 10004, 10th floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. I agree to submit a yearly progress report.

Principal Investigator Signature Date

**PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES**

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Ms Michelle McCabe

| | |
|---|---|
| **CLEARANCE CERTIFICATE** | **M120994** |
| **PROJECT** | Microsatellite Instability Assessment in Colorectal carcinoma Tumour Biopsies: Mismarch Repair Immunohistochemistry versus Microsatellite PCR |
| **INVESTIGATORS** | Ms Michelle McCabe, |
| **DEPARTMENT** | School of Anatomical pathology |
| **DATE CONSIDERED** | Ad hoc |
| **DECISION OF THE COMMITTEE*** | Approved unconditionally |

**Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.**

**DATE** 28/09/2012 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor :

**DECLARATION OF INVESTIGATOR(S)**
To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

*PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...*

APPENDIX B:

Plagiarism declaration for written work

I Michelle McCabe, as a postgraduate student registered for a PhD at the University of the Witwatersrand declare that this thesis is my own written work, except where I have stated otherwise and indicated by referencing the work, thoughts and ideas of others



Signature ……..