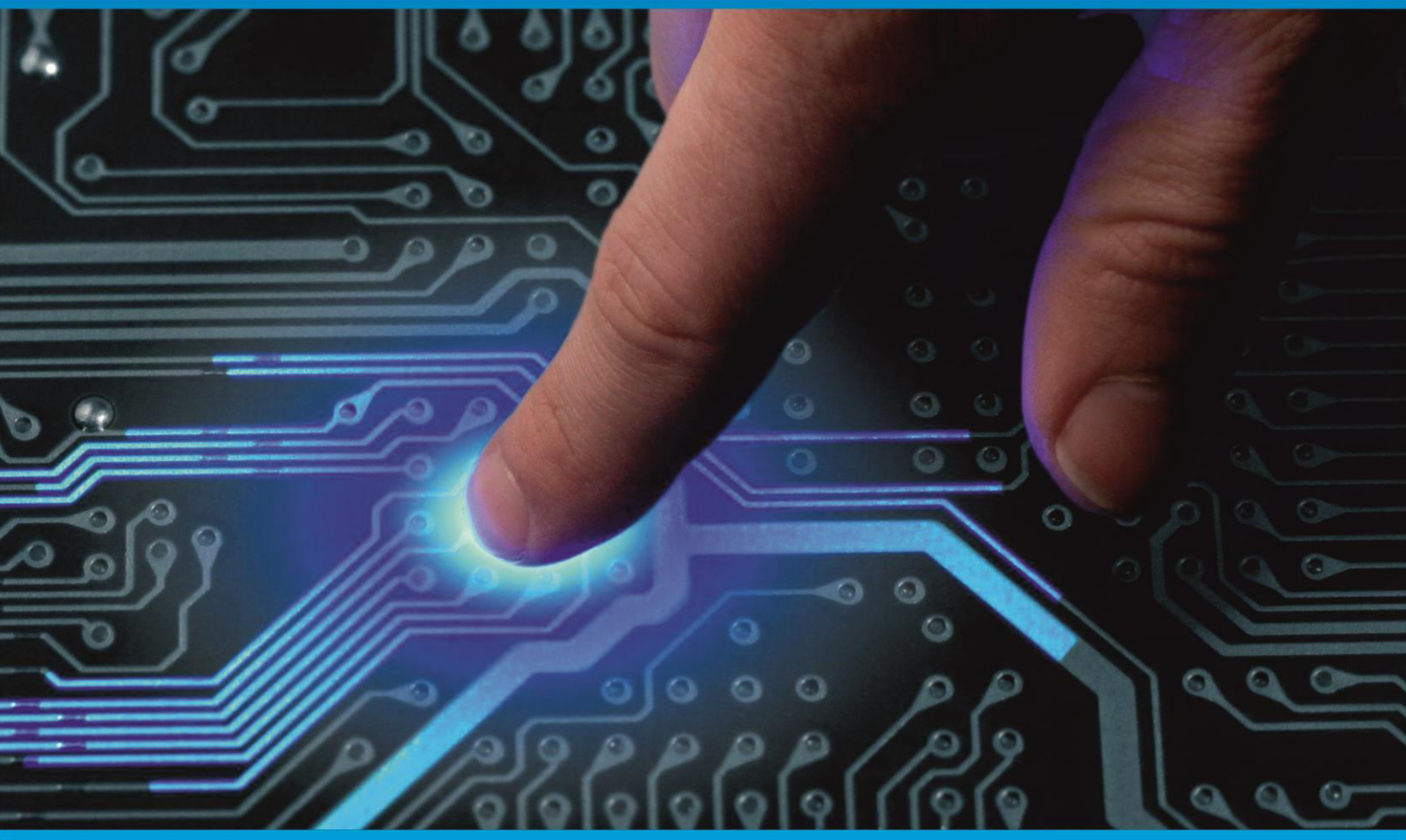
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905088 |

A Review on Risk Prediction Models for Colorectal Cancer

Gayatri Laware, Ranjit Gawande

PG Student, Dept. of Computer Engineering, Matoshri College of Engineering & Research Centre, Eklhare, Nashik, India

Assistant Professor, Dept. of Computer Engineering, Matoshri College of Engineering & Research Centre, Eklhare, Nashik, India

ABSTRACT: Predictive risk models are essential to identify high-risk individuals with disease, and then can provide them with individual clinical treatment, screening and interventions to reduce the burden of disease. They can also be used for research purposes when trying to identify new risk factors for disease. In this article, we will review the risk prediction model developed for colon cancer and evaluate its applicability, advantages and disadvantages. We will also discuss the factors that need to be considered in the future development and improvement of colorectal cancer risk prediction models. Exist No model can adequately cover the known risk factors of colorectal cancer in order to screen the entire risk range, so a new comprehensive model is needed.

KEYWORDS: Colorectal Cancer, Risk Factor, Prediction, Risk evaluation Model

I. INTRODUCTION

Colon cancer (CRC) is one of the most commonly diagnosed cancers in the world. In 2008, more than one million new cases were diagnosed (9.8% of global cancer diagnoses) and 600,000 deaths (8.1% of all cancer deaths worldwide) (1). There may be widespread risks in the population (2).Although this is partly due to individual differences in exposure to environmental risk factors, theoretically speaking, due to potential family risk factors (we call "family profile risks"), the risks will be very different (3-7) .The relationship between family history as a risk factor and the main family risk status indicated: (i) In the lowest quartile, the risk of CRC changed 20 times (mean value 1).The lifetime risk of CRC is 25%), and the highest quartile of the family risk profile (average risk is 25%; Reference 4); (ii) 90% of all CRCs occur in people above the average family risk profile.

A. Risk factors for CRC

Many risk factors have been involved in the development of CRC. [9] Family history is a known risk factor. [10] Studies have shown that people with first-degree relatives (father, offspring, brother or sister) have an average risk of CRC compared with people without a family history [10-12], people with 3 or more affected first-degree relatives [10] have a three-fold increase in association. However, the above statistics are only average values. There is a family history, and the age of the high-risk group depends on the age when the sick relatives are diagnosed (the earlier the age at diagnosis, the more likely the high-risk group is to be born) And the kinship between them (the greater the number and/or the closer the kinship, the more likely they are to be at risk;[10].

At best, only half of the risk of familial CRC can be explained by known high-risk genetic mutations [11]. Although a large amount of research funds have been used to find other genes that are susceptible to CRC, these genes have not been confirmed. On the other hand, genome-wide association studies have identified at least 15 universal markers (Single Nucleotide Polymorphisms, SNP) for genetic susceptibility. These markers have a slight increase in the risk of developing CRC (or homozygous) Related. Small allele). Compared with non-operator rankings0.80 to 1.70; (Lit. 12-14) The reasons for these weak associations, not to mention residual familial risks, may include relatively rare genetic variants and/or other risk factors shared by relatives.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905088 |

B. Predicting Risk of CRC

Risk prediction models are important because they can be used to (i) reduce the burden of disease by taking preventive measures against the most vulnerable groups; (ii) identify those with the most genetic predisposition, and then perform genetic testing (iii) Improve the effectiveness of observational research in identifying new risk factors for disease (provided that this predictive risk model can highly predict high-risk individuals with CRC). Observational research data (including case-control studies and cohort studies) are often used to develop predictive risk models.

1. Reducing the burden of disease.

Healthcare managers and clinicians rely on risk assessments to decide who to test for CRC. The variables (including new predictors, such as genomic data) are that they tend to be more accurate than the clinical stage alone. The predictive model identifies the people most likely to benefit from CRC screening or other preventive interventions. Most CRC is caused by polyps, which are a known precancerous stage that may appear many years before symptoms (such as rectal bleeding, bowel movements, or anemia) appear.

2.Identification of new risk factors for developing CRC

In observational studies on the risk of CRC, traditionally, the environment, lifestyle, and incidence of genetic risk factors between affected people and healthy people (case-control design) or CRC among exposed and untreated people have been compared. Incidence rate (cohort design). In most of these studies, participants were selected regardless of their family history of CRC, therefore, most of the population did not have a family history of the disease. Therefore, the results of this study cannot be generalized to individuals with a family history of the disease. Since the analysis is usually applied to family history, the observed associations only represent the average association in the broad family risk profile [8]. It is known which risk factors play specific roles in the category of household risk status, or whether the association differs depending on the household risk status.

Predictive risk models can be used as tools to identify emerging risk factors for diseases. One possible approach is to classify individuals into high-risk and low-risk individuals, and then compare cases and controls in those low-risk and high-risk categories for a given exposure level (assuming that they are not used for risk classification). This method has tremendous statistical power, can determine other risk factors for diseases, and can provide clues about the interaction between genes and the environment.

II. EXISTING RISK PREDICTION MODELS FOR CRC

Based on known genetic and environmental risk factors, a risk prediction model for viral CRC has been developed. The purpose of this article is to summarize these models and the research that evaluates them in terms of applicability, strengths, and weaknesses. CRC prediction model, including risk factors and/or cancer susceptibility genes related to the individual, environment and lifestyle. To focus our review, we excluded models that predict CRC staging based on patient and tumor characteristics (egCai et al.;, and models that predict colorectal tumor types (benign or malignant). Clinical symptoms (e.g. Brothers and colleagues; [15], a model that predicts the incidence of CRC based on the growth rate of polyps after the first polypectomy (eg Wilson and Lightwood), predicts symptomatic CRC based on the intestinal symptoms assessed by the therapist Model predicts patients (S. Selvachandran et al.; and models that are based only on the first cancer and gene signature.

A. Non-genetic models

Previously developed models for predicting CRC risk use a regression model scoring system that includes family medical history, lifestyle, and environmental risk factors. These models are easy to implement, and non-genetic risk factors can be easily incorporated. However, it is difficult to provide family information. Medical history, including number of relatives, age of healthy relatives, age at diagnosis other than first-generation relatives, and related risk factors between family members (including genetic and non-genetic factors). CRC risk Although models containing a family history of colon cancer (16) or CRC (17,18) (binary (yes/no)) represent the average risk of low-risk groups, they cannot provide accurate information for high-risk groups for the following reasons Risk assessment: There may be a strong family history or the risk of known genetic mutations.

Another important question about the history of colonoscopy is the need for screening or diagnostic colonoscopy. A negative colonoscopy indicates a reduced risk of disease (compared to those who did not have a colonoscopy; reference [19], and a positive colonoscopy indicates an increased risk of disease (signs of susceptibility). More adenomas; [19], and if possible, reduce the risk of illness Polypectomy can remove precancerous adenomas [20]. It is not clear how

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905088 |

such information should be included in the risk prediction model. This will be important as colonoscopy is increasingly used as a diagnostic or screening test.

B. Genetic models

This model has several limitations: (i) Only the family history of CRC of the second-degree relatives is used (Ii) Does not contain the PMS2-MMR gene, which accounts for 15% of MMR mutations, although since the risk estimate of mutation frequency in this range is reliable, this may have little impact; (iii) Does not contain any environmental risk factors; (iv)) Cannot predict the second primary cancer risk of affected individuals. None of the above models (summarized in Table 1) can fully explain its complexity For example, it is known that high-risk mutations in CRC susceptibility genes account for up to half of the family risk Genetic association studies have identified SNPs at 15 loci, which are respectively associated with a slight increase in the risk of CRC. The model suggests that if the remaining family risk is due to similar genetic effects, there may be hundreds of genetic changes to the risk of CRC. Although a cancer prediction model with a multi-gene component has been developed, this component represents the influence of multiple genetic variants, but each variant has little effect on the risk of breast cancer and prostate cancer. A comparable tool has not been developed for CRC.

III. EVALUATION OF RISK PREDICTION MODELS

Before a risk prediction model can be recommended as a useful tool for personal decision-making in a clinical setting, it must be tested in a population that is not related to the population used to create the model. The prediction models are as follows:

(i) The calibration (or reliability) assesses the ability of the model to predict the number of events to be predicted (CRC in this case). Use goodness of fit or x2 for comparison expected number of events with the observed number of events.

(ii) The discrimination (or precision) measurement model uses the adjusted statistic (statistic c) corresponding to the range below the "performance" feature to distinguish between people who are more likely to develop a disease and people who are less likely to develop a disease Ability. Recipient. The curve or the classification index only (NRI) represents the probability that the improved model will correctly reclassify people (ie, Will increase the risk score of people with CRC or decrease the risk score of people without CRC) minus the possibility of incorrect reclassification (that is, reduce the risk score of people with CRC or increase the risk score) does not belong to someone A person who does not belong to that person;[21]

(iii) The accuracy assessment model is the most likely and most likely probability of getting the disease in some people, and the person's usefulness in predicting the risk of the disease. Sensitivity and specificity as well as positive and negative predictive values are important indicators of this test.

(iv) Utility, that is, the ability of the model to be supplemented by the people it targets (such as doctors, patients, the general population, and policy makers). From user surveys or interviews.

A. Harvard Cancer Risk Index

This model, including 2 qualitative studies (22,23). Although this model is well calibrated for women, it overestimates the CRC in low-risk men (24). By evaluating non-professional risk understanding, risk perception and result interpretation, it is found that this method is very popular among users. [24]A computerized tool that provides absolute personal risk and relative CRC estimates (Emmons et al. [24], qualitatively evaluated the colorectal cancer risk assessment and communication tools studied by Harvard, to ensure the accuracy of risk perception as a concern Point and user satisfaction, the conclusion is that the tool can be used to correct the misunderstanding of personal risk. At the beginning of the study, the risk perception was inaccurate, and more than half of the participants in the intervention group corrected their risk perception at the last time Compared with the test, the control group is only 12%.

B. Imperial's model

Imperiale and colleagues [25] validated their model using an independent data set from the same source of the developed model. 1,031 men and women with no bowel symptoms underwent colonoscopy and found corresponding actual risk of progressive lesions in the proximal colon (tubular adenomas larger than 1 cm, polyps with hair histology or severe dysplasia or cancer) And what was predicted. The model evaluates low, medium, and high risks, although it

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905088 |

shows moderate discrimination statistic 0.74 [24]. However, because different models and validation data sets are from the same population, the validity of the model for different populations is unknown.

C. Freeman's model

Park and colleagues [26] evaluated the Freeman model using prospective data from 260,000 people, with an average follow-up time of 7 years. They found that Freeman's model was well calibrated for men and women and most risk factor categories. However, the model overestimates the risk of people with a family history of CRC, and the number of CRC is expected to be 35% higher than that of men with CRC.A relative with CRC has 42% more men with a history of polyps. The model underestimated the risk of men and women who had previously been screened but did not have polyps, ranging from 33% to 0.61, respectively.

D. Ma's model

Ma and colleagues (27) verified their own model based on the study of the Japan Public Health Center. They found that their model underestimated the incidence of colon cancer by 19% (95% CI, 3-37%), but the protocol is beneficial for rectal methods. Cancer and CRC in general. This model underestimates CRC cases with 4 environmental categories (age, body mass index, drinking, and smoking).

IV. FUTURE PERSPECTIVE

Existing CRC risk models are limited in terms of built-in risk factors and their effectiveness. Based on known disease risk factors and hypothetical but residual risk factors, a more complex model is needed. Consider developing a comprehensive model to predict CRC risk.

A. Environmental factors

The role of physical characteristics and environmental exposure in the risk of CRC may depend on the presence of genetic mutations [26]. The previous model did not consider the interaction of genes with the environment or the interaction of genes. Explaining these interactions is not easy. It is observed that the strength of the association between environmental risk factors and CRC in people with a family history is different from that in people randomly selected from the population (30-31), which is consistent with the existence of a genetic environment. However, to date, several studies have directly compared the strength of associations between wearers and non-wearers. We have found that for carriers and non-carriers, BMI is similar to the risk of CRC in early adulthood. Fruit consumption fiber intake (32), and smoking (32-33) did not directly compare the strength of the association between carriers and non-carriers.

B. Need for ethnicity-specific risk models

Most CRC risk prediction models are developed using data mainly from white people, and therefore may not be applicable to other racial/ethnic groups. These populations need further research. The difference is up to 10 times (1, 2) internationally. Although there may be some genetic differences in CRC risk, studies have shown that most of these differences are due to: With the increase in the incidence of CRC among immigrants, the differences in environmental risk factors approach the host country within one to two generations.

C. Extracolonic cancers

Hereditary CRC syndromes are rarely limited to the colon and rectum. According to reports, the risk of several extracolonic cancers is increased: uterine cancer, gastric cancer, ovarian cancer, ureteral cancer, renal pelvis cancer, brain cancer, small bowel cancer, and hepatobiliary tract cancer (34); familial adenomatous polyposis Brain, thyroid, and liver cancer (35); uterine and gastric cancer (36) in carriers of monoallelic MUTYH mutations; and duodenal, bladder, skin, and ovarian cancers in carriers of biallelic MUTYH mutations. Even in families that are not known to have high-risk mutations, people with a family history of extra colon cancer have a higher risk of CRC. Therefore, the impact of these other familial cancers on the risk of CRC should be included in the risk model.

V. CONCLUSION AND FUTURE WORK

In order to accurately determine the risk of a person suffering from CRC, it is important to establish a genetic-based predictive model. The model should consider family history and environmental risk factors of several generations, and also consider the simultaneous exposure of different MMR genes and other genes. High disease susceptibility.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905088 |

Including the influence of MUTYH and low permeability genes, as well as residual genetic risk factors or other family risk factors. Only such a model can provide an accurate estimate of future cancer risk and predict the high-risk gene mutation status of everyone in the entire CRC risk range. By implementing the predictive model as a web application, the predictive model should be easy to evaluate and use by clinicians, genetic counsellors, and the general public.

REFERENCES

1. Ferlay J, Shin H, Bray F, Forman D, Mathers C, Parkin D. GLOBOCAN 2008, cancer incidence and mortality worldwide. Lyon, France: International Agency for Research on Cancer; 2010.
2. Jemal A, Center MM, DeSantis C, Ward EM. Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiol Biomarkers Prev2010; 19:1893–9.
3. Aalen OO. Modelling the influence of risk factors on familial aggregation of disease. Biometrics 1991;47:933–45.
4. Hopper JL, Carlin JB. Familial aggregation of a disease consequent upon correlation between relatives in a risk factor measured on a continuous scale. Am J Epidemiol 1992; 136:1138–47.
5. Khoury MJ, Beaty TH, Liang KY. Can familial aggregation of disease be explained by familial aggregation of environmental risk factors? Am J Epidemiol 1988;127:674–83.
6. Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 2002;31:33–6.
7. Peto J, Mack TM. High constant incidence in twins and other relatives of women with breast cancer. Nat Genet 2000;26:411–4.
8. Hopper JL. Disease-specific prospective family study cohorts enriched for familial risk. EpidemiolPerspectInnov 2011;8:2.
9. Potter JD. Risk factors for colon neoplasia–epidemiology and biology. Eur J Cancer 1995:31A:1033–8.
10. Taylor DP, Burt RW, Williams MS, Haug PJ, Cannon-Albright LA. Population-based family history-specific risks for colorectal cancer: a constellation approach. Gastroenterology 2010;138:877–85.
11. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. Am J Gastroenterol 2001;96:2992– 3003.
12. Baglietto L, Jenkins MA, Severi G, Giles GG, Bishop DT, Boyle P, et al. Measures of familial aggregation depend on definition of family history: meta-analysis for colorectal cancer. J Clin Epidemiol 2006; 59:114–24.
13. Aaltonen L, Johns L, Jarvinen H, Mecklin J-P, Houlston R. Explaining € the familial colorectal cancer risk associated with mismatch repair (MMR)-deficient and MMR-stable tumors. Clin Cancer Res 2007;13: 356–61.
14. Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, Spain S, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet 2007;39:984–8.
15. Brazer SR, Pancotto FS, Long TT III, Harrell FE Jr, Lee KL, Tyor MP, et al. Using ordinal logistic regression to estimate the likelihood of colorectal neoplasia. J Clin Epidemiol 1991;44:1263–70.
16. Colditz G, Atwood K, Emmons K, Monson R, Willett W, Trichopoulos D, et al. Harvard report on cancer prevention volume 4: Harvard Cancer Risk Index. Cancer Causes Control 2000; 11:477–88.
17. Freedman AN, Slattery ML, Ballard-Barbash R, Willis G, Cann BJ, Pee D, et al. Colorectal cancer risk prediction tool for white men and women without known susceptibility. J Clin Oncol 2009;27:686–93.
18. Wei EK, Colditz GA, Giovannucci EL, Fuchs CS, Rosner BA. Cumulative risk of colon cancer up to age 70 years by risk factor status using data from the Nurses' Health Study. Am J Epidemiol 2009;170: 863–72.
19. Schoen RE. Surveillance after positive and negative colonoscopy examinations: issues, yields, and use. Am J Gastroenterol 2003;98: 1237–46.
20. Brenner H, Chang-Claude J, Seiler CM, Rickert A, Hoffmeister M. Protection from colorectal cancer after colonoscopy: a population based, case-control study. Ann Intern Med 2011;154:22–3.
21. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008; 27:157–72; discussion 207–12.
22. Kim DJ, Rockhill B, Colditz GA. Validation of the Harvard Cancer Risk Index: a prediction tool for individual cancer risk. J Clin Epidemiol 2004;57:332–40.
23. Emmons KM, Koch-Weser S, Atwood K, Conboy L, Rudd R, Colditz G. A qualitative evaluation of the Harvard Cancer Risk Index. J Health Commun 1999;4:181–93.
24. Emmons KM, Wong M, Puleo E, Weinstein N, Fletcher R, Colditz G. Tailored computer-based cancer risk communication: correcting colorectal cancer risk perception. J Health Commun 2004;9: 127–41.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905088 |

25. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Using riskfor advanced proximal colonic neoplasia to tailor endoscopic screening for colorectal cancer. Ann Intern Med 2003;139:959–65.
26. Park Y, Freedman AN, Gail MH, Pee D, Hollenbeck A, Schatzkin A, et al. Validation of a colorectal cancer risk prediction model among white patients age 50 years and older. J Clin Oncol 2009;27: 694–8.
27. Ma E, Sasazuki S, Iwasaki M, Sawada N, Inoue M. 10-Year risk of colorectal cancer: development and validation of a prediction model in mid.
28. Campbell PT, Cotterchio M, Dicks E, Parfrey P, Gallinger S, McLaughlin JR. Excess body weight and colorectal cancer risk in Canada: associations in subgroups of clinically defined familial risk of cancer. Cancer Epidemiol Biomarkers Prev 2007;16:1735–44.
29. Campbell P, Newcomb P, Gallinger S, Cotterchio M, McLaughlin J. Exogenous hormones and colorectal cancer risk in Canada: associations stratified by clinically defined familial risk of cancer. Cancer Causes Control 2007;18:723–33.
30. Slattery ML, Levin TR, Ma K, Goldgar D, Holubkov R, Edwards S. Family history and colorectal cancer: predictors of risk. Cancer Causes Control 2003;14:879–87.
31. Win AK, Dowty JG, English DR, Campbell PT, Young JP, Winship I, et al. Body mass index in early adulthood and colorectal cancer risk for carriers and non-carriers of germline mutations in DNA mismatch repair genes. Br J Cancer 2011;105:162–9.
32. Watson P, Ashwathnarayan R, Lynch HT, Roy HK. Tobacco use and increased colorectal cancer risk in patients with hereditary nonpolyposis colorectal cancer (Lynch syndrome). Arch Intern Med 2004;164: 2429–31.
33. Diergaarde B, Braam H, Vasen HF, Nagengast FM, van Muijen GN, Kok FJ, et al. Environmental factors and colorectal tumor risk in individuals with hereditary nonpolyposis colorectal cancer. Clin Gastroenterol Hepatol 2007;5:736–42.e
34. 101. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, et al. Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 2004;96:261–8.
35. Vasen HF, Moslein G, Alonso A, Aretz S, Bernstein I, Bertario L, et al. Guidelines for the clinical management of familial adenomatous polyposis (FAP). Gut 2008;57:704–13.
36. Win AK, Cleary SP, Dowty JG, Baron JA, Young JP, Buchanan DD, et al. Cancer risks for monoallelic MUTYH mutation carriers with a family history of colorectal cancer. Int J Cancer 2011;129: 2256–62.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details