LATEST TECHNIQUE

A Novel Approach in Elucidating the Interaction of Ecology, Biology and Social (Eco-bio-social) Factors with Serum Metabolites and Its Pathway among Polyps and Colorectal Cancer: A Study Protocol

ABU BAKAR MF[1], MOHAMMED NAWI A[1], CHIN SF[2], MAKPOL S[3]

[1]Department of Public Health Medicine, [3]Department of Biochemistry, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Yaacob Latiff, Bandar Tun Razak, 56000 Cheras, Kuala Lumpur, Malaysia
[2]UKM Medical Molecular Biology Institute (UMBI), Universiti Kebangsaan Malaysia, Jalan Yaacob Latiff, Bandar Tun Razak, 56000 Cheras, Kuala Lumpur, Malaysia

ABSTRAK

Peningkatan jumlah insiden dan kadar kematian disebabkan oleh kanser kolorektal (KKR) telah menyebabkan peningkatan beban kepada sistem penjagaan kesihatan negara-negara yang terlibat. Ini telah memberi kesan kepada komponen ekonomi yang penting sebagai satu beban yang ditanggung oleh sesebuah negara termasuklah penjagaan perubatan dan kos bukan perubatan yang telah menyebabkan kerugian produktiviti di kalangan pesakit dan penjaga. Diagnosis awal adalah penting untuk membantu dalam mengurangkan kadar kematian disebabkan oleh KKR dan meningkatkan keberkesanan rawatan penyakit. Selain faktor genetik, ketidakseimbangan gaya hidup dan faktor persekitaran juga dilaporkan sebagai punca utama KKR. Walau bagaimanapun, kebanyakan kajian terdahulu hanya memberi tumpuan kepada satu faktor pada satu masa di mana ia telah menghadkan pemahaman tentang bagaimana faktor-faktor ini boleh dimanipulasi untuk mengkaji penanda-bio penyakit. Oleh itu, kajian ini membantu dalam membina protokol baru untuk memberikan gambaran yang lebih baik mengenai kesan ekologi, biologi, dan faktor sosial yang mungkin mempengaruhi tahap pembezaan serum metabolit bagi polip dan KKR yang berpotensi digunakan untuk membezakan pesakit dengan polip atau KKR. Penemuan kajian ini berpotensi untuk mengenalpasti penanda-bio dari serum yang boleh digunakan dalam pembangunan alat saringan baru untuk meningkatkan kadar saringan awal polip dan KKR.

Address for correspondence and reprint requests: Assoc. Prof. Dr Azmawati Mohammed Nawi. Department of Public Health Medicine, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Yaacob Latiff, Bandar Tun Razak, 56000 Cheras, Kuala Lumpur, Malaysia. Tel: +6019-3131340 Email: azmawati@ppukm.ukm.edu.my

Kata kunci: Alat saringan; ekobiososial faktor; kanser kolorektal; penanda-bio serum; polip

ABSTRACT

The increasing number of colorectal cancer (CRC) incidences and mortality rates have caused an increasing burden on the healthcare system of the involved countries. This will impact the important economic components as a burden including direct medical care and non-medical costs and have caused productivity losses among patients and caregivers. Early diagnosis is important to help in reducing mortality rates of CRC and enhancing disease treatments. Besides genetic pathways, lifestyle imbalance and environmental factors also are reported as the main causes of CRC. However, most of the previous studies only focused on one factor at a time which limit the understanding of how these factors could be manipulated in the disease biomarkers finding. Hence, this study helps to build a novel protocol to provide a better insight into the effects of the ecology, biology, and social factors on the differential level of a polyp and CRC serum metabolites that could potentially be used to distinguish patients with polyp or CRC. The finding of the study could serve as a potential biomarker that could be used in the development of new screening tools to enhance early screening of polyps and CRC.

Keywords: Serum biomarker; Colorectal cancer; Polyps; Eco-bio-social factors; screening tool

INTRODUCTION

Colorectal cancer (CRC) has been reported as one of the major causes of cancer-related death worldwide. According to Siegel et al. (2020) among adults aged less than 50 years, CRC case rates increased by 22% from 2000 to 2013, driven solely by tumors in the distal colon and rectum. Similar to incidence patterns, CRC death rates decreased by 34% among individuals aged more than 50 years from 2000 to 2014 but increased by 13% in those aged less than 50 years (Siegel et al., 2020). According to World Health Organisation (WHO) (2020) in the cancer country profile report, CRC is the third leading cancer in Malaysia with an incidence rate of 14% and mortality rate of 12.9%. A recent study reported that accumulated evidence indicates that CRC is a genetically heterogeneous and complicated disease caused by abnormalities in gene expression and its structure. It is now generally accepted that CRC mainly develops through two different genetic pathways, one is the pathway involving microsatellite instability (MSI), while the other is the chromosomal instability (CIN) pathway (Zhou et al. 2015).

Besides genetic pathways, CRC was

also reported to be caused by a lifestyle imbalance and eating habits. For example, people with obesity or who did not include fruits and vegetables in their dietary routine have a high risk to get diagnosed with CRC compared to the person who might carry the genetic traits but has a healthy lifestyle (Sawicki et al. 2021). Moreover, according to Soffian et al. (2021) identified factors linked with CRC cluster can be further classified into ecology (health care accessibility, urbanicity, dirty streets, tree coverage), biology (age, sex, ethnicity, overweight and obesity, daily consumption of milk and fruit, family history of CRC, level of metabolites), and social determinants (median income level, smoking status, health cost, employment status, housing violations, and domestic violence). However, none of the experimental studies that combine these three main factors all together has been conducted previously.

CRC is mostly asymptomatic until it progresses to advanced stages with low overall survival rates. Thus, early detection is essential to reducing incidence and mortality rates (Tepus & Yau 2020). It has been reported that the progression against CRC can be accelerated by increasing initiation of screening, while late diagnosis usually will lead to poor progression and a low survival rate (Bray et al. 2018). Current CRC diagnosis methods are mainly depending on colonoscopy and sigmoidoscopy which are invasive procedures, have potential risk for complications, and require high cost and skilled operators (Vega et al. 2015). Therefore, non-invasive methods

have been invented including fecal immunochemical tests and fecal occult blood tests. However, these methods have low accuracy and sensitivity in detecting the early stages of CRC (Gonzalez-Pons & Cruz-Correa 2015; Nishiumi et al. 2012; Tepus & Yau 2020). Hence, the need for screening methods that are non-invasive, specific, and accurate for early identification of CRC has spurred many researchers to turn to the use of molecular techniques such as genomics, proteomics, and more recently, metabolomics, to identify serum biomarkers (Hashim et al. 2021).

In terms of metabolites study, metabolomics profiling can be conducted to measure the level of each detected metabolite in CRC biological matrices. It is downstream of genomics, transcriptomics, and proteomics (Hashim et al. 2021). The alterations at the metabolomics level not only reflect the changes at the genomics and proteomics levels but also are influenced by environmental factors. Differences in the levels of metabolites between diseased and normal states are used to identify altered metabolic pathways in CRC patients (Hashim et al. 2021). The advancement of technology nowadays has allowed researchers to develop a measurable metabolomic profiling pipeline (Long et al. 2017). For CRC, metabolomic profiling has become one of the potential methods to determine disease etiology and biomarkers. However, the association between the differential level of metabolites in CRC patients and other Eco-bio-social factors is rarely reported.

An effort must be made to understand the interaction between the ecological, biological, and social (Eco-bio-social/EBS) determinant factors, and their influences on the level of metabolites in CRC patients which can lead to the design and implementation of improved and sustainable CRC case control. Hence, this study aimed to identify Eco-bio-social factors and understand their association with the differential level of signature metabolites related to CRC initiation and progression, which may eventually be used in CRC diagnosis and prognosis, and reduce the mortality rates among CRC patients.

PREVIOUS GAP THIS STUDY WANTS TO FILL IN

Numerous studies have reported an ecology, or biology, or social factors associated with polyp and CRC (Rawla et al. 2019; Sawicki et al. 2021; Takahashi & Nakao 2021), but none of them has investigated a combination of all driving factors (Eco-bio-social factors) for CRC and polyp in one study. Moreover, most CRC serum metabolomics studies, such as Nishiumi et al. (2012), Tan et al. (2013) and Hashim et al. (2021), focused on the differentiation between normal healthy patients and CRC patients. However, there is a limited study that has included colorectal polyp in their study design. In addition, Hashim et al. 2021 also suggested that the differences of the list of serum metabolites associated with CRC reported from previous studies might be due to the environmental and genetics factors.

Hence, this study will be conducted to determine the main and shared of all Eco-bio-social factors between polyp and CRC, identify serum biomarkers associated with polyp and CRC, and understand the association or effects of those factors with differential level of serum metabolites biomarker for polyp and CRC in which can be potentially use in the development of new screening tool.

STUDY OBJECTIVES

In this study, the ecology, biology, and social factors among control, polyp, and CRC patients will be compared followed by identification of a signature metabolites associated with polyp and CRC using global metabolomics profiling. The pathway analysis will be conducted to determine the altered metabolic pathways in polyp and CRC patients followed by the joint effects analysis between a panel of signature metabolites of polyp and CRC with ecology, biology, and social factors to investigate the association between these factors with signature metabolites associated with polyps and CRC.

EXPERIMENTAL DESIGN

The study will be conducted at the Endoscopy centre, Hospital Canselor Tunku Mukhriz (HCTM), Universiti Kebangsaan Malaysia, from April 2023 until April 2024, using a case-control study design. The cases will include all patients whom undergo a colonoscopy and are diagnosed with polyp or CRC based on the histology results at HCTM. Meanwhile, all patients

undergo a colonoscopy are confirmed to have healthy normal colons based on colonoscopy results at HCTM will be acted as a control group. All participants will be selected based on the inclusion and exclusion criteria in Table 1 with the ratio of 1: 1:1.

Figure 1 shows the flowchart of the methodology design. This study will be divided into two phases; Phase 1: Questionnaire distribution and blood sampling, and Phase 2: Computational and statistical analysis. Figure 1 shows the flowchart of the methodology design for this study. For phase I, questionnaire distribution and blood sampling will be conducted where all potential participants will be selected from the name list of the patients who undergo colonoscopy at HCTM. Based on the colonoscopy result, the patients will be categorised into three groups: normal, polyp, and CRC. The questionnaire will be given to the selected participants who agreed to participate in this study followed by blood sampling for metabolomic profiling. The blood samples are then

will be centrifuged at 4°C and 4500 rpm for 10 minutes to obtain serum blood plasma. The serum samples will be stored at -80°C until further analysis.

In Phase II, the serum samples will be analysed using global metabolomic profiling to identify the list of metabolites present in the samples. The samples will be run through Ultra High-Performance Liquid Chromatography-Mass Spectrometry (UHPLC-MS) and will be prepared in accordance with the standard guidelines (Hashim et al. 2021; Long et al. 2017). The data collected will be quantitatively analysed to obtain the list of potential metabolites and will be submitted to MetaboAnalyst 5.0 (Pang et al. 2021) for metabolic pathway analysis. From the analysis, the top 5 signature metabolites will be selected based on their significant interaction in polyp and CRC pathways. Next, these five signature metabolites will be validated using UHPLC-MS via targeted metabolomic profiling. The same serum samples will be used in the validation process. Finally, the

TABLE 1: Sampling criteria

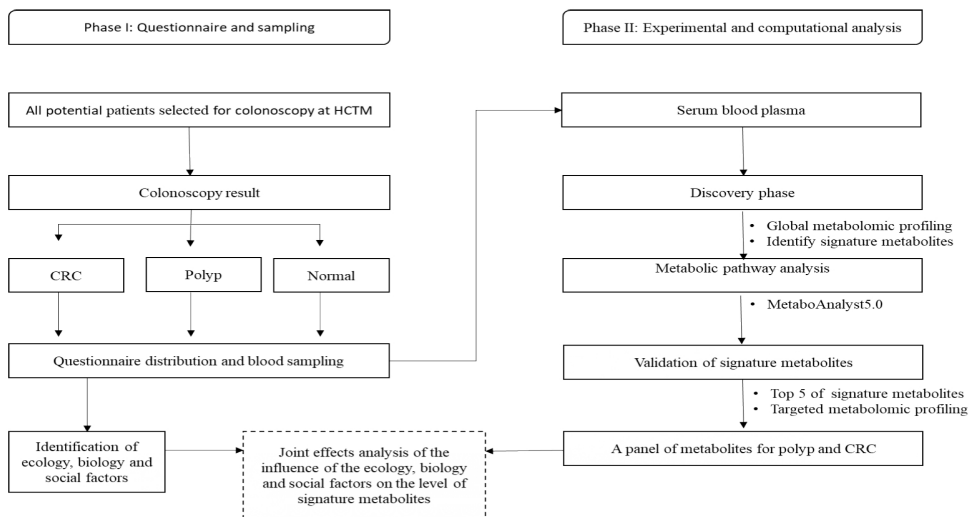|	Control	Colon polyp	CRC
Research method		Case-control	
Location		Endoscopy centre, HCTM	
Inclusion criteria	-Malaysian -Age 18 and above -Have the ability to read and converse in Malay and English. -Do not diagnose with colon polyp and CRC based on colonoscopy histology result	-Malaysian -Age 18 and above -Have the ability to read and converse in Malay and English. - Diagnose with colon polyp based on colonoscopy and histology result	-Malaysian -Age 18 and above -Have the ability to read and converse in Malay and English. -Diagnose with CRC based on colonoscopy and histology result
Exclusion criteria	Patients with inflammatory bowel disease	Patients with inflammatory bowel disease	CRC patients with more than one cancer
Ratio	1	1	1

FIGURE 1: Flowchart of the methodology design

joint-effects analysis will be carried out using STATA software (StataCorp. 2021) to identify the association between the Eco biosocial factors and a panel of metabolites of polyp and CRC. The summary of experimental metabolomic analysis in phase II was further explained in Figure 2.

PATIENT AND PUBLIC INVOLVEMENT

All the patients and clients on the name list who undergo colonoscopy at HCTM from April 2023 until April 2024. There are three sampling units i.e. CRC patients, colon polyp patients, and healthy colon patients. The CRC patients are patients who have undergo colonoscopy at HCTM from April 2023 until April 2024, who have been diagnosed histologically with CRC, undergo treatment, and meet the inclusion and exclusion criteria.
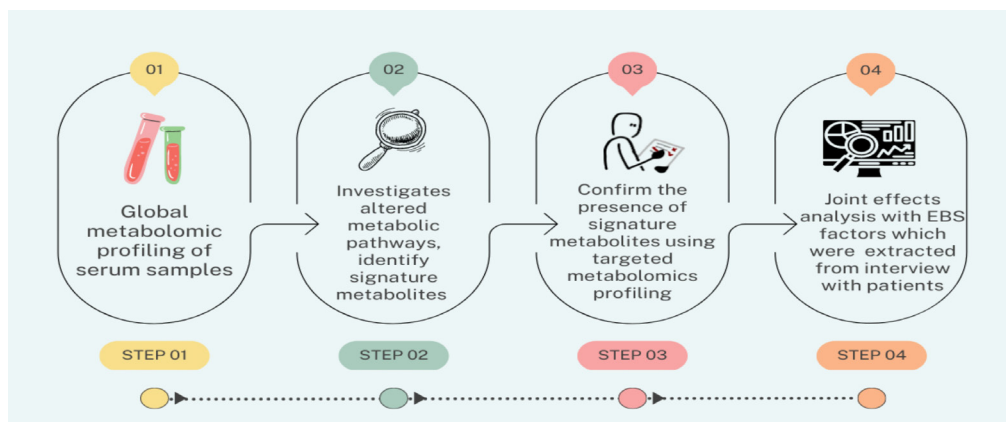

FIGURE 2: Summary of experimental analysis

Meanwhile, colon polyp patients are patients who have undergo colonoscopy at HCTM from April 2023 until April 2024 who have been certified histologically to have a polyp and meet the inclusion and exclusion criteria. As for the control, patients who have undergo colonoscopy at HCTM from April 2023 until April 2024 who have been certified normal and meet the inclusion and exclusion criteria will be considered as control patients.

SAMPLE SIZE

The calculation of the sample size was carried out using Software PS Power and Sample Size Calculations Version 3.1.6 (Dupont & Plummer 1998) as shown in Table 2. From the calculations, the required sample is estimated from 88 to 857. Due to the time and financial constraints, the sample size required for this study will be set at 62 for each healthy control group, colon polyps, and CRC. In total, 223 samples are required with an increase of 20% to include dropped out or incomplete respondents. The same sample will be used in the validation set.

QUESTIONNAIRE

The Malay questionnaire will be handed to the selected participants. It was divided into 4 sections: Section A (sociodemographic and socioeconomic), section B (biological and clinical factors), section C (social and lifestyle factors), and section D (awareness level of CRC). The details of the questionnaire are as follow:

Section A (sociodemographic and socioeconomic):

In this section, there are questions regarding the background of the respondents including sociodemographic information on age, gender, race and socio-economic information on monthly income, occupation, education level, and address.

Section B (Biological and clinical)

For biological factors, the questions about the history of cancer in the family will be asked to the respondents. The questions included the question about family cancer history, family relationships, and the type of cancer they had. These questions were adapted from the questionnaire "Family History Questionnaire (FHQ)" used in a study conducted in the Netherlands (Kessels et al. 2017). This questionnaire has been translated into Malay language and validated by Norsa'adah et al. (2020). The permission to use this questionnaire has been obtained. Meanwhile, for clinical factors, questions on comorbidity, histological type and CRC stage will be asked (for CRC group only).

Section C (lifestyle and social factors)

The questions in this section are related to lifestyle such as smoking status, alcohol consumption, body mass index (BMI), physical activity and food intake patterns. This information is adapted from studies on CRC and

TABLE 2: Sample size calculation

Objectives	Factors	Author/country	Value	Software	N	Additional 20% for unresponsive or incomplete sample
Compare the ecology, biology, and social factors among control, polyp and CRC patients	Smoking	Chang et al. (2021)/ Canada	Smoking status with young onset CRC P0=0.32 P1=0.21 M=1.0	Software PS Power and Sample Size Calculations Version 3.1.6 (Dupont & Plummer 1990)	252	605
	BMI (≥25.0)	Cho et al. (2019)/ Korea	Association of BMI ≥25 with CRC P0=0.36 P1=0.49 M=2.0		168	403
	Obesity	Ulaganathan et al. (2018)/ Malaysia	Obesity association with CRC P0=0.54 P1=0.67 M=1.0		221	530
	Family history of CRC	Alqahtani et al. (2020), Arab Saudi	Association of family history with CRC as CRC driving factor P0=0.07 P1=0.16 M=1.0		196	470
	Metabolites					
Identify signature metabolites associated with polyp and CRC using global metabolomics profiling and gene expression analysis respectively	Hypoxanthine	Long et al. (2017), USA	Level of hypoxanthine in CRC vs. polyp δ:0.34 σ:0.67 M:1.0	Software PS Power and Sample Size Calculations Version 3.1.6 (Dupont & Plummer 1990)	62	205
	D-mannose		Level of D-mannose in CRC vs. polyp δ:0.28 σ:0.57 M:1.0		66	271
	Xanthine		Level of xanthine in CRC vs. polyp δ:0.33 σ:0.42 M:1.0		26	62
	Glutamate		Level of glutamate in CRC vs polyp δ:0.30 σ:0.81 M:1.0		113	271

N: sample size; P0: probability of exposure in control; P1: probability of exposure in cases ; M: ratio of control to experimental subjects; δ: difference in population means ; σ: standard deviation

risks from lifestyle (Nawi et al. 2020).

Section D (awareness level of CRC)

There are ten questions related to the awareness of CRC risk factors in this section. The questions were adopted from The Bowel/Colorectal Cancer Awareness Measure (CAM) which was developed by the University college of London and Cancer Research UK, in 2008 and was translated in into the Malay language by (Su et al., 2013). Responses were given on the Likert scale which ranged from 'strongly disagree, 'disagree', 'uncertain, agree' and 'strongly agree'. Each answer "agree" or "strongly agree" will get one score; and the maximum score is 10 marks. High scores indicate a high awareness level of CRC.

From this questionnaire, the ecology factors will be identified from the address given by the participants including the data such as the coordinated position of fast-food restaurants, pedestrian walks, green spaces, and distance from hospitals. Meanwhile, the biology factors will be determined from section B (family history of cancer, underlaying illnesses, differential level of metabolites, etc). Moreover, differential level of serum metabolites among participants will be further analysed using global metabolomics profiling. Lastly, the social factors will be determined from the data extracted from section C and section D.

DATA ANALYSIS

IBM SPSS software (IBM Corp. 2021) will be used to conduct univariate and bivariate analysis for chi-square and ANOVA tests to determine the main and shared factors associated with colorectal cancer between healthy control, colon polyps, and CRC. The significant value will be set at the p-value of <0.05. For social factors, the data will be collected from the questionnaire, while the data for biological factors will be collected from the results of the metabolomic profiling. Furthermore, the data such as the coordinated position of fast-food restaurants, pedestrian walks, green spaces, and hospitals will be sourced from Google Maps and the calculation of the distance (in kilometers) between the address obtained in section A and respected places will be collected for ecology factors. These data are then will be tested using multinomial logistic regression analysis to identify if they are the driving factors for CRC and polyp. The result will be shown in the form of an odds ratio with a confidence interval of 95%.

Next, the data from the UHPLC-MS analysis will be collected and data normalisation will be performed using MetaboAnalyst 5.0 for statistical analysis (Pang et al. 2021). Chemometric and univariate analyses will be performed using MetaboAnalyst 5.0. Principal Component Analysis (PCA) will be selected as the tool of chemometric analysis to visualise the difference and outliers of the groupings in the form of a score plot. Univariate analysis using a volcano plot will be performed following PCA scoring by using the parameter of fold-change threshold > 2 and p-value less than 0.1

as suggested by Fan et al. (2020). List of metabolic features generated from a volcano plot will be further identified using m/zcloud and Chemspider included in Xcalibur™ version 3.1 (Termo Fisher Scientific, USA), CEU Mass Mediator (CEUMM)108, Human Metabolome Database (HMDB)72, and METLIN (Montenegro-Burke et al. 2020).

In addition, the list of potential metabolites will be submitted to the MetaboAnalyst 5.0 (Pang et al. 2021) for pathway analysis which will be used to construct a network and determine the relations between the identified metabolites and their corresponding genes and pathways. From the network constructed, the top 5 signature metabolites that show significant importance in polyp and CRC pathway will be selected for a validation process. Finally, the joint-effects analysis will be carried out using STATA software (StataCorp. 2021). Each signature metabolite differentiated between CRC vs Normal, Polyps vs. Normal, and Polyps vs. CRC will be assessed as variables using multinomial logistic regression analysis to investigate the influence of the Eco-bio-social factors on their expression level in CRC, colon polyp, and control group. The odds ratio (OR) and the corresponding 95% confidence interval (CI) will be calculated with multivariate logistic regression models, which will be adjusted for respected factors such as age, ethnicity, gender, and smoking status. In addition, the interaction for identified factors (for example smoking, alcohol, and BMI) that might influence CRC progression will be added to the multivariate models to determine the interaction between the signature metabolites level and each identified Eco-bio-social factor in modulating CRC risk. All statistical tests will be set as 2-sided with the significance set at $P < 0.05$.

EXPECTED RESULTS

The findings of this study could suggest a panel of serum metabolites biomarker that can be used to distinguish patients with polyp and CRC. Moreover, a positive association between eco-bio-social factors with polyp and CRC serum metabolites biomarker could provide a better understanding on how environmental, social, and biological factors could affect a differential level of signature metabolites among polyp and CRC population.

ACKNOWLEDGEMENT

FUNDING

INSTITUTIONAL REVIEW BOARD STATEMENT

The work has been approved by the Universiti Kebangsaan Malaysia (UKM) ethics committee (UKM PPI/111/8/JEP-2022-603).

REFERENCES

Alqahtani, M.Z., Mohammed, A.G., Alsamghan, A.S., Bharti, R.K., Alsharm, A.A., Alshahrani, M.T., Alzahrani, M.A., Ayed, A.A.N., Alsaleem, M.A., Ghazwani, E.Y. (2020). Risk factors of colorectal cancer among Saudi Population: Case-control study. *J Family Med Prim Care* **9**(9): 5035-40.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **68**(6): 394-424.

Chang, V.C., Cotterchio, M., De, P., Tinmouth, J. 2021. Risk factors for early-onset colorectal cancer: a population-based case–control study in Ontario, Canada. *Cancer Causes Control* **32**(10): 1063-83.

Cho, Y.A., Lee, J., Oh, J.H., Chang, H.J., Sohn, D.K., Shin, A., Kim, J. 2019. Genetic risk score, combined lifestyle factors and risk of colorectal cancer. *Cancer Res Treat* **51**(3): 1033-40.

Dupont, W.D., Plummer Jr, W.D. 1998. Power and sample size calculations for studies involving linear regression. *Control Clin Trials* **19**(6): 589-601.

Fan, S., Shahid, M., Jin, P., Asher, A., Kim, J. 2020. Identification of metabolic alterations in breast cancer using mass spectrometry-based metabolomic analysis. *Metabolites* **10**(4): 170.

Gonzalez-Pons, M., Cruz-Correa, M. 2015. Colorectal cancer biomarkers: Where are we now?. *BioMed Res Int* 2015: 149014.

Hashim, N.A.A., Ab-Rahim, S., Ngah, W.Z.W., Nathan, S., Ab Mutalib, N.S., Sagap, I., Jamal, A.R.A., Mazlan, M. 2021. Global metabolomics profiling of colorectal cancer in Malaysian patients. *BioImpacts* **11**(1): 33.

IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp.

Kessels, K., Eisinger, J.D., Letteboer, T.G., Offerhaus, G.J.A., Siersema, P.D., Moons, L.M. 2017. Sending family history questionnaires to patients before a colonoscopy improves genetic counseling for hereditary colorectal cancer. *J Dig Dis* **18**(6): 343-8.

Long, Y., Sanchez-Espiridion, B., Lin, M., White, L., Mishra, L., Raju, G.S., Kopetz, S., Eng, C., Hildebrandt, M.A., Chang, D.W., Ye, Y. 2017. Global and targeted serum metabolic profiling of colorectal cancer progression. *Cancer* **123**(20): 4066-74.

Montenegro-Burke, J.R., Guijas, C., Siuzdak, G. 2020. METLIN: A tandem mass spectral library of standards. *Methods Mol Biol* 2104: 149-63.

Nawi, A.M., Chin, S.F., Mazlan, L., Jamal, R. 2020. Delineating colorectal cancer distribution, interaction, and risk prediction by environmental risk factors and serum trace elements. *Sci Rep* **10**(1): 18670.

Nishiumi, S., Kobayashi, T., Ikeda, A., Yoshie, T., Kibi, M., Izumi, Y., Okuno, T., Hayashi, N., Kawano, S., Takenawa, T., Azuma, T. 2012. A novel serum metabolomics-based diagnostic approach for colorectal cancer. *PloS ONE* **7**(7): e40459.

Norsa'adah, B., Aminu, A.R., Zaidi, Z. 2020. Validity and reliability of questionnaire on knowledge, attitude and dietary practices related to colorectal cancer. *Malays J Med Sci* **27**(1): 115-23.

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D.A., Chang, L., Barrette, M., Gauthier, C., Jacques, P.É., Li, S., Xia, J. 2021. MetaboAnalyst 5.0: Narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res* **49**(W1): W388-W396.

Rawla, P., Sunkara, T., Barsouk, A. 2019. Epidemiology of colorectal cancer: Incidence, mortality, survival, and risk factors. *Prz Gastroenterol* **14**(2): 89-103.

Sawicki, T., Ruszkowska, M., Danielewicz, A., Nied wiedzka, E., Ar ukowicz, T., Przybyłowicz, K.E. 2021. A review of colorectal cancer in terms of epidemiology, risk factors, development, symptoms and diagnosis. *Cancers* **13**(9): 2025.

Siegel, R.L., Miller, K.D., Goding Sauer, A., Fedewa, S.A., Butterly, L.F., Anderson, J.C., Cercek, A., Smith, R.A., Jemal, A. 2020. Colorectal cancer statistics, 2020. *CA Cancer J Clin* **70**(3): 145-64.

Soffian, S.S.S., Nawi, A.M., Hod, R., Chan, H.K., Hassan, M.R.A. 2021. Area-level determinants in colorectal cancer spatial clustering studies: a systematic review. *Int J Environ Res Public Health* **18**(19): 10486.

StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC.

Su, T.T., Goh, J.Y., Tan, J., Muhaimah, A.R., Pigeneswaren, Y., Khairun, N.S., Normazidah, A.W., Tharisini, D.K., Majid, H.A., 2013. Level of colorectal cancer awareness: A cross sectional exploratory study among multi-ethnic rural population in Malaysia. *BMC Cancer* **13**(1): 1-8.

Takahashi, N., Nakao, M. 2021. Social-life factors associated with participation in screening and further assessment of colorectal cancer: A nationwide ecological study in Japanese municipalities. *SSM Popul Health* **15**: 100839.

Tan, B., Qiu, Y., Zou, X., Chen, T., Xie, G., Cheng, Y., Dong, T., Zhao, L., Feng, B., Hu, X., Xu, L.X., Zhao, A., Zhang, M., Cai, G., Cai, S., Zhou, Z., Zheng, M., Zhang, Y., Jia, W. 2013. Metabonomics identifies serum metabolite markers of colorectal cancer. *J Proteome Res* **12**(6): 3000-9.

Tepus, M., Yau, T.O. 2020. Non-invasive colorectal cancer screening: An overview. *Gastrointest*

Tumors **7**(3): 62-73.

Ulaganathan, V., Kandiah, M., Shariff, Z.M. 2018. A case–control study on the association of abdominal obesity and hypercholesterolemia with the risk of colorectal cancer. *J Carcinog* **17**: 4.

Vega, P., Valentin, F., Cubiella, J. 2015. Colorectal cancer diagnosis: Pitfalls and opportunities. *World J Gastrointest Oncol* **7**(12): 422.

World Health Organization (WHO). Cancer Country Profile Report 2020. International Agency for Research on Cancer (IARC). Available at https://gco.iarc.fr/ [Accessed 27 May 2022]

World Medical Association. WMA Declaration of Helsinki: ethical principles for medical research involving human subjects, 2018. Available: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [Accessed 6 February 2023].

Zhou, X., Xu, X., Wang, J., Lin, J., Chen, W. 2015. Identifying miRNA/mRNA negative regulation pairs in colorectal cancer. *Sci Rep* **5**(1): 12995.