



Review

# Area-Level Determinants in Colorectal Cancer Spatial Clustering Studies: A Systematic Review

Sharifah Saffinas Syed Soffian <sup>1</sup>, Azmawati Mohammed Nawi <sup>1,\*</sup> , Rozita Hod <sup>1</sup> , Huan-Keat Chan <sup>2</sup> and Muhammad Radzi Abu Hassan <sup>2</sup>

<sup>1</sup> Department of Community Health, Faculty of Medicine, Universiti Kebangsaan Malaysia, Kuala Lumpur 56000, Malaysia; p108896@siswa.ukm.edu.my (S.S.S.); rozita.hod@ppukm.ukm.edu.my (R.H.)

<sup>2</sup> Clinical Research Center, Sultanah Bahiyah Hospital, Alor Setar 05400, Malaysia; huankeat123@yahoo.com (H.-K.C.); drradzi91@yahoo.co.uk (M.R.A.H.)

\* Correspondence: azmawati@ppukm.ukm.edu.my

**Abstract:** The increasing pattern of colorectal cancer (CRC) in specific geographic region, compounded by interaction of multifactorial determinants, showed the tendency to cluster. The review aimed to identify and synthesize available evidence on clustering patterns of CRC incidence, specifically related to the associated determinants. Articles were systematically searched from four databases, Scopus, Web of Science, PubMed, and EBSCOHost. The approach for identification of the final articles follows PRISMA guidelines. Selected full-text articles were published between 2016 and 2021 of English language and spatial studies focusing on CRC cluster identification. Articles of systematic reviews, conference proceedings, book chapters, and reports were excluded. Of the final 12 articles, data on the spatial statistics used and associated factors were extracted. Identified factors linked with CRC cluster were 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), and social determinants (median income level, smoking status, health cost, employment status, housing violations, and domestic violence). Future spatial studies that incorporate physical environment related to CRC cluster and the potential interaction between the ecology, biology and social determinants are warranted to provide more insights to the complex mechanism of CRC cluster pattern.

**Keywords:** colorectal cancer; risk factors; cluster; geographical information system; ecobiosocial



**Citation:** Soffian, S.S.S.; Nawi, A.M.; Hod, R.; Chan, H.-K.; Hassan, M.R.A. Area-Level Determinants in Colorectal Cancer Spatial Clustering Studies: A Systematic Review. *Int. J. Environ. Res. Public Health* **2021**, *18*, 10486. <https://doi.org/10.3390/ijerph181910486>

Academic Editor: Marc Saez

Received: 19 August 2021

Accepted: 3 October 2021

Published: 6 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cancer is one of the most important causes of mortality and morbidity around the globe. It is the third cause of death after cardiovascular diseases and road traffic injuries as reported by the Global Burden of Disease Study 2017 [1]. Among the types of cancers, colorectal cancer (CRC) accounted as the third most common cancer in men and the second highest in women [2]. The disease represented the loss of 15,800,000 disability-adjusted life years in 2013, 56% of which in middle- and low-income countries and 44% in the industrialized countries [3]. While the trend of colorectal cancer had shifted to the left in the western regions, the new cases diagnosed among the young and elderly age group Asians are increasing [4,5].

CRC is multifactorial by nature. No single hazardous factor is plausibly related to CRC, but individual factors such as sex, age, and family history, lifestyle behaviors including alcohol consumption, high intake of red meat and processed meat, low fruit and vegetable intake, high-fat diet, and physical inactivity were massively studied [6–8]. Few researchers postulated on the specific gene–environment interaction likely to cause CRC, on an individual basis [9,10]. Although education on the risk factors for CRC have been

continuously delivered to the public, there are still evidences of huge disparities in CRC incidence across different location descriptively [11,12].

While many factors were found related to CRC, generally they can be classified into modifiable and non-modifiable. Among the established modifiable risk factors include obesity, westernized diet, physical inactivity, and low fiber intake [7,13–15]. Meanwhile, the non-modifiable risk factors include hereditary, age, gender, and ethnicity [16–18]. Numerous studies in the literature have focused on colorectal cancer because of its high incidence and mortality and that it is closely related to individual lifestyle (modifiable risk factors), indirectly the tendency to cluster. People living in the same neighborhood tend to have similar lifestyle and share many cluster-inducing factors that grow substantial public concern over locally elevated CRC incidence [19,20]. Despite of the knowledge on modifiable and non-modifiable risk factors, less is known about interaction between multiple risk factors that may occur within a small geographical area. Thus, there has been an expansion of studies that explore the relationship between simultaneous risk factors collectively contributing to the potential of CRC cluster within a local area.

According to the classical theory of epidemiological triad, disease transmission can be explained through host, agent, and environment factors and the interaction that overlapped with each other. The triad model had shown successful intervention to curb the spread of infectious disease. The recognition of multiple risk factors related to CRC proposes the implementation of ecobiosocial approach to explain the relationship and interaction between the environment, biological, and social. Historically, the concept of ecobiosocial was essential in the field of vector-borne diseases, whereby integrated vector management actions and planning were first developed [21]. In 2016, an obesity framework based on the ecobiosocial concept was attempted to further elucidate the significance of environmental influence that shape the unhealthy choice of foods, compounded with social and genetic factors leading to the issue of obesity [22]. As CRC is a chronic disease, requiring more than a decade of potential exposure to the multiple risk factors, it may result from the complex interaction of ecobiosocial factors. The ecological was referred to as “individual activity” and “activity environment” [22] that disrupt the energy imbalance promoting obesity, one of the CRC risk factors. This includes any factors that promote or inhibit the physical activity of an individual within a local setting such as the availability of recreational park, green areas, safe walkability, or accessible public transportation as provided in many major cities. The biological component features the predisposing genetic factors, age, sex, ethnicity, and obesity depicted the susceptibility of an individual towards developing CRC. Notwithstanding that, the social factors define the behaviors and attitudes are likely shaped by the cultural importance of eating pattern, westernized diet, socioeconomic status, smoking habit, and lack of health-seeking behavior.

Spatial cluster detection is an important tool in cancer surveillance to identify areas of elevated risk and to generate subsequent hypotheses about cancer etiology [23,24]. Provided that the existence of known established risk factors of an area, spatial cluster analysis may predict the future trend of cancer locally and inform control strategies. A spatial disease cluster is defined as an area with an unusually higher disease incidence rate [25]. However, the term has been vaguely used to refer to a population-based cancer epidemiology due to the complex interaction between multiple factors believed to contribute to such event. Cancer cluster identification is heavily dependent on the accuracy of methodological design used to estimate the local relative risk as compared to the control [26,27]. Besides, spatial analysis of CRC incidence may provide a new knowledge on the relationships between external risk factors and people lifestyle with CRC burden across communities. This will enable policymakers to develop tailored intervention to areas where the CRC risk is greater. Therefore, the review aimed to identify and synthesize available evidence on clustering patterns of CRC incidence, specifically related to the associated determinants.

## 2. Materials and Methods

The systematic review was conducted in compliant to the PRISMA or Preferred Reporting Items for Systematic Reviews and Meta-Analyses review protocol [28]. PRISMA aims to guide researchers source the appropriate information at high level of accuracy. Based on the protocol, the authors initiated the systematic literature review by formulating appropriate research questions. The authors performed systematic searching that consists of identification, screening, and eligibility. The authors then proceed to appraise the quality of the selected articles using the Appraisal tool for Cross-Sectional Studies (AXIS tool) [29] to ensure the quality of the articles included. Upon completion, the authors read through in detail all the articles for data extraction and analysis.

### 2.1. Formulation of the Research Question

The research question was formulated based on the PICO concept; a tool often used to assist authors in developing suitable research questions for the review. It consists of Population or Problem, Interest, and Context/Outcome [30]. Based on this concept, the authors have included the three main aspects in the review: colorectal cancer (Population), spatial cluster (Interest), and determinants (Context/Outcome), which led the authors to the main research question “What are the determinants commonly linked to colorectal cancer cluster in spatial studies?”.

### 2.2. Systematic Searching Strategies

The systematic searching strategy preceded by the identification, screening, and eligibility stages (Figure 1).

### 2.3. Identification

Identification stage involved enrichment of the keywords through the utilization of synonyms and their variation to be used during article searching in the databases. The search string was developed and computed using Boolean operators and phrase searching as shown in Table 1. The systematic literature search was conducted between 24 and 27 May 2021, which involved four primary databases: Scopus, Web of Science, PubMed, and EBSCOHost, resulted in the retrieval of 3134 records. These four databases were selected because of their availability and accessibility to our institution. There were 45 duplicate records found and removed. The records were exported from the databases and arranged for screening in an Excel sheet.

**Table 1.** Keyword search used in the identification process.

Database	Search String
Scopus	TITLE-ABS-KEY (“colorectal cancer” OR “colorectal neoplasm * <sup>1</sup> ” OR “colorectal tumor *” OR “colorectal carcinoma” OR “large bowel cancer”) AND (“cluster analysis” OR “spatial analysis” OR “geographical information system” OR “geographic distribution” OR “incidence distribution” OR “demography”) AND (“risk factor *” OR “cancer risk” OR “determinant *”)
Web of Science	TS = (“colorectal cancer” OR “colorectal neoplasm *” OR “colorectal tumor *” OR “colorectal carcinoma” OR “large bowel cancer”) AND (“cluster analysis” OR “spatial analysis” OR “geographical information system” OR “geographic distribution” OR “incidence distribution” OR “demography”) AND (“risk factor *” OR “cancer risk” OR “determinant *”)
PubMed	(“colorectal cancer” OR “colorectal neoplasm *” OR “colorectal tumor *” OR “colorectal carcinoma” OR “large bowel cancer”) AND (“cluster analysis” OR “spatial analysis” OR “geographical information system” OR “geographic distribution” OR “incidence distribution” OR “demography”) AND (“risk factor *” OR “cancer risk” OR “determinant *”)

Table 1. Cont.

Database	Search String
EBSCOHost	((“colorectal cancer” OR “colorectal neoplasm *” OR “colorectal tumor *” OR “colorectal carcinoma” OR “large bowel cancer”) AND (“cluster analysis” OR “spatial analysis” OR “geographical information system” OR “geographic distribution” OR “incidence distribution” OR “demography”)) AND (“risk factor *” OR “cancer risk” OR “determinant *”))

<sup>1</sup> The symbol \* was used in the search strategy as truncation and wildcard function to increase variability of selected keywords.

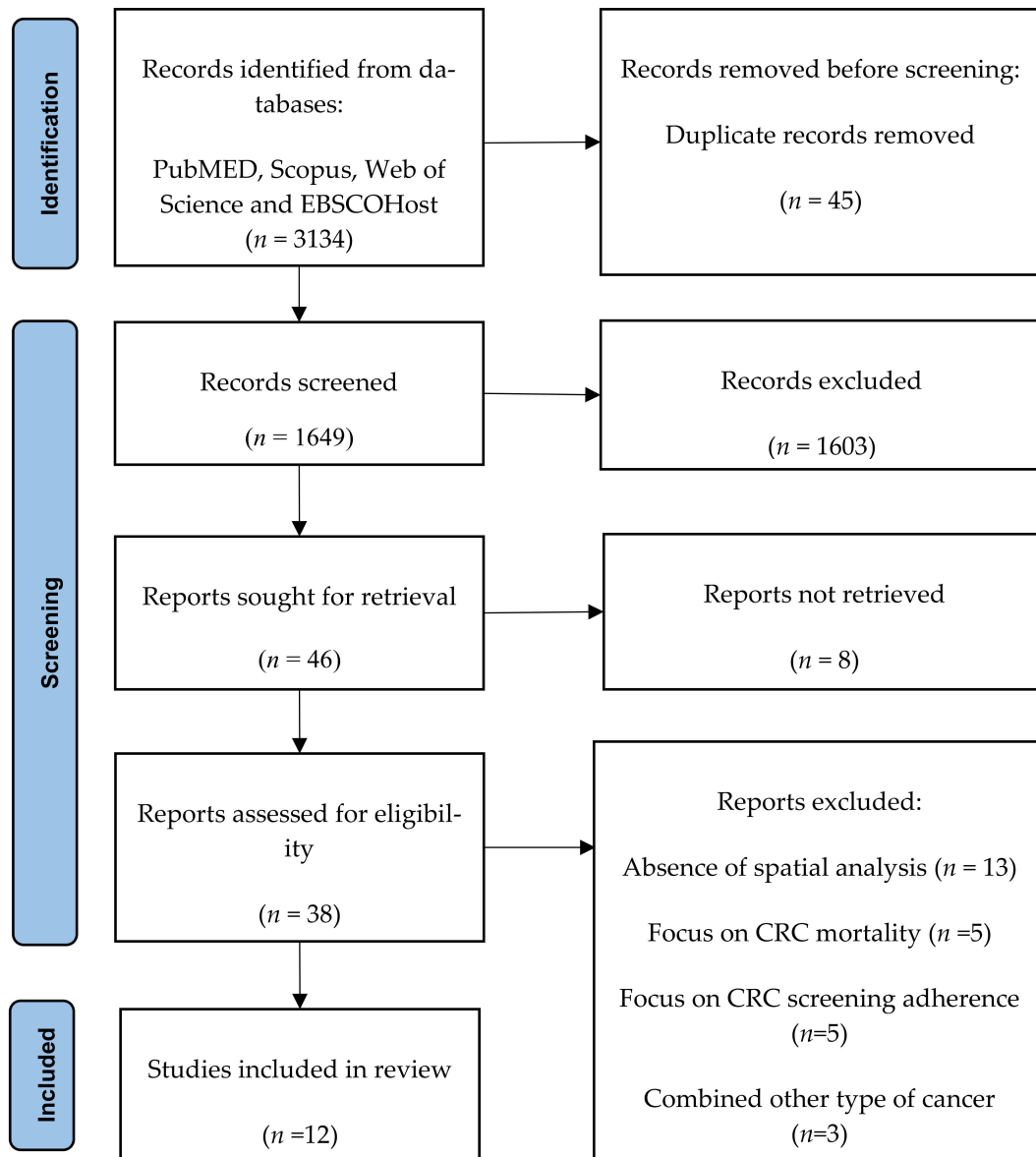


Figure 1. PRISMA flow diagram.

#### 2.4. Screening

The title and abstract of each article were examined for relevance and screened based on specific criteria by the authors. The inclusion criteria for article selection were: (1) published between 2016 and 2021, (2) full original article, (3) written in English, (4) study focused at identifying colorectal cancer clusters or interrelations between two or more of the clusters. The duration of published articles screened was determined based on the

recent development and dynamicity of Geographic Information System (GIS) software. Articles of systematic review, conference proceeding, book chapter, and reports were excluded. Any disagreement on article selection was resolved via discussion. The screening process had excluded 1603 articles, while the remaining 46 articles proceeded for retrieval of full text for eligibility.

### 2.5. Eligibility

There were 38 full text articles successfully retrieved for eligibility. The authors reviewed all full text articles and recorded the reason for the article exclusion. A total of 26 articles were excluded due to the absence of spatial analysis ( $n = 13$ ), the articles focus on CRC mortality ( $n = 5$ ), focus on CRC screening adherence ( $n = 5$ ) and article that combined other type of cancer ( $n = 3$ ). The remaining articles were resumed with the quality appraisal process.

### 2.6. Quality Appraisal

The articles selected from the eligibility process must be further examined for risk of bias assessment to ensure the quality of the study [31]. Study quality was assessed using the appraisal tool for observational and cross-sectional studies (AXIS tool) as shown in Table 1. The scale is designed for non-experimental research and includes 20 items measuring each aspect of study quality [29]. Each study was assessed for potential risk of biases through key domain areas of study design, sample size justification, target population, sampling frame, sample selection, measurement validity and reliability, methodology limitations and discussion. Two authors conducted the assessments independently. Any disagreement between the two authors was resolved through discussion until consensus met; when necessary, a third reviewer was consulted. A total of 12 articles were included in the final stage.

The result for quality assessment is presented in Table 2. The total number of “yes” were recorded for every study as the tool guide does not specify the standardize scoring measure. The mean total quality score was 15.4 (range 14–16). Of the 12 studies finalized, quality assessment using AXIS revealed all the included studies had clear study objectives and employed appropriate study design with respect to their objectives. Similarly, all 12 studies clearly defined the target population with an appropriate sampling frame. Only three studies addressed and categorized non responders [32–34]. The risk factor and outcome variables measured were appropriate to the aims of each study and were correctly measured. All studies clearly explained the statistical significance used and sufficiently described the methodology to enable them to be repeated. One study [35] inadequately described the basic data. Meanwhile, none of the studies reported information on the nonresponders, possibly due to the nature of ecological analysis used in spatial studies. All studies provided information on the methodological limitations. Five studies did not state any information regarding ethical approval or consent of participants [33,36–39].

### 2.7. Data Abstraction and Analysis

The authors independently extracted information from the included studies on author’s names, year, country, study objectives, study designs, sample, statistical test used, factors associated with cluster, additional findings, and limitations. Discrepancies were resolved through discussion between the two authors. Following that, the authors systematically synthesize the outcomes based on the use of words and text to explain the findings according to patterns, themes, consistencies, inconsistencies, and relationship within the extracted data [40].

Data abstracted from all studies were recorded in an appropriate matrix table (Table 3). The authors reviewed the matrix tables for consistencies and inconsistencies to generate themes and findings for the review. Similar or related information were grouped together as one characteristic and the technique repeated to form reasonable findings for interpretation. The authors discussed respective thoughts associated with findings until the point of agreement was reached upon adjustment of the generated findings.

**Table 2.** Quality assessment of included studies using AXIS tool.

Articles	Introduction				Methods							
	Author (Year), Country	Were the aims/objectives of the study clear?	Was the study design appropriate for the stated aim(s)?	Was the sample size justified?	Was the target/reference population clearly defined?	Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Were measures undertaken to address and categorise non-responders?	Were the risk factor and outcome variables measured appropriate to the aims of the study?	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted, or published previously?	Is it clear what was used to determine statistical significance and/or precision estimates?	Were the methods (including statistical methods) sufficiently described to enable them to be repeated?
Liu et al. (2016), United States	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mansori et al. (2018), Iran	Y	y	y	y	y	y	CT	Y	Y	Y	Y	Y
Mansori et al. (2019), Iran	Y	Y	Y	Y	Y	Y	CT	Y	Y	Y	Y	Y
Pakzad et al. (2016), Iran	Y	Y	Y	Y	Y	Y	CT	Y	Y	Y	Y	Y
Pourhoseingholi et al. (2020), Iran	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Goungounga et al. (2016), France	Y	Y	Y	Y	Y	Y	CT	Y	Y	Y	Y	Y
Roquette et al. (2019), Portugal	Y	Y	Y	Y	Y	Y	CT	Y	Y	Y	Y	Y
Torres et al. (2018), United States	Y	Y	Y	Y	Y	Y	CT	Y	Y	Y	Y	Y
Halimi et al. (2019), Iran	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Harminder et al. (2017), Canada	Y	Y	Y	Y	Y	Y	CT	Y	Y	Y	Y	Y
Kuo et al. (2019), United States	Y	Y	CT	Y	Y	Y	CT	Y	Y	Y	Y	Y
Goshayeshi et al. (2019), Iran	Y	Y	Y	Y	Y	Y	CT	Y	Y	Y	Y	Y

Table 2. Cont.

Articles	Results			Discussion				Other		Total Recorded “Yes”
	Were the basic data adequately described?	Does the response rate raise concerns about non-response bias?	If appropriate, was information about non-responders described?	Were the results internally consistent?	Were the results for the analyses described in the methods, presented?	Were the authors’ discussions and conclusions justified by the results?	Were the limitations of the study discussed?	Were there any funding sources or conflicts of interest that may affect the authors’ interpretation of the results?	Was ethical approval or consent of participants attained?	
Liu et al. (2016), United States	Y	N	N	CT	Y	Y	Y	N	Y	16
Mansori et al. (2018), Iran	Y	CT	CT	CT	Y	Y	Y	N	CT	14
Mansori et al. (2019), Iran	Y	CT	CT	CT	Y	Y	Y	N	CT	15
Pakzad et al. (2016), Iran	Y	N	CT	Y	Y	Y	Y	CT	CT	15
Pourhoseingholi et al. (2020), Iran	Y	N	CT	Y	Y	Y	Y	N	CT	16
Goungounga et al. (2016), France	Y	N	CT	Y	Y	Y	Y	N	Y	16
Roquette et al. (2019), Portugal	N	N	CT	Y	Y	Y	Y	N	Y	15
Torres et al. (2018), United States	Y	N	N	Y	Y	Y	Y	N	Y	16
Halimi et al. (2019), Iran	Y	N	N	CT	Y	Y	Y	N	Y	16
Harminder et al. (2017), Canada	Y	N	N	CT	Y	Y	Y	N	Y	15
Kuo et al. (2019), United States	Y	N	CT	Y	Y	Y	Y	N	Y	15
Goshayeshi et al. (2019), Iran	Y	N	CT	Y	Y	Y	Y	N	CT	16

Table 3. Characteristics of included studies.

No.	Author (Year), Country	Objective	Study Design	Sample	Statistical Test	Findings	Factors Associated	Determinants	Limitation
1	Liu et al. (2016), United States	To determine the association between median household income and the risk of developing colorectal cancer in Texas	Retrospective cohort	155,534 men and women with colorectal cancer in 1995–2011	Getis-Ord G, Poisson regression model (ArcGIS v10.1)	Higher median household income, measured from the third to the highest income quintile ranging between \$38,040 and \$80,876, was associated with decreased risk of colorectal cancer in Texas. The Hispanics showed higher incidence rate of CRC compared to the Non-Hispanic Whites and Blacks at all time period, with slight decrease trend across higher median household income.	County median household income level (Median household incomes in 254 counties classified into five quintiles) Ethnicity (Non-Hispanic Whites, Non-Hispanic Blacks, Hispanics)	Social Biology	Income variable was the median household income measured at county level, thus potential ecological fallacy. Median household income alone may not be a good proxy for SES. Lack of accuracy in population estimates may lead to biased calculation for CRC incidence
2	Mansori et al. (2018), Iran	To determine the factors associated with the spatial distribution of the CRC incidence in the neighborhoods of Tehran, Iran	Ecological	2815 new cases of CRC from 2008 to 2011	The Moran Index, BYM model (using OpenBUGS version 3.2.3, ArcGIS 10.3, GeoDa)	There was spatial autocorrelation at the level of the neighborhoods. Significant association was observed between women head of household, living in a rental house, no daily milk consumption in the household and higher household health expenditures against higher risk of CRC respectively.	Socioeconomic variables (SES): unemployed people aged 15 years and above, educated women (university level of education) aged 17 years old and above, women head of household, households without a car, those living in a rental house, households with income below the poverty line, people without insurance coverage. Risk factors: Households without daily fruit consumption, households without daily milk consumption, overweight people aged 15 years and above, smoking households. Health costs: Household health expenditures, expenditures on diagnoses, expenditures on medicine, expenditures on hospitals, expenditures on medical visits	Social Biology	Ecological fallacy. Edge effect, referring to border neighborhoods affected by size of adjacent regions. Misclassification in geocoding due to incomplete postal addresses.
3	Mansori et al. (2019), Iran	To determine effective factors on geographic distribution of the CRC incidence in Iran	Ecological	2815 new cases of CRC from 2008 to 2011	Geographically Weighted Poisson Regression Model (using GWR 4, Stata 14, ArcGIS 10.3)	The spatial variability was observed with more frequent utilization of health services as indicated by the household health expenditures (Median Incidence Risk Ratios (IRR): 1.39), the cost of diagnosis of the disease (Median IRR: 1.03), the cost of household medicine (Median IRR: 1.05), the cost of hospital admission (Median IRR: 1.09) and the cost of medical visits (Median IRR: 1.27)	Socioeconomic variables (SES) include employment status, education Daily fruit consumption, daily milk consumption, overweight, smoking Health costs variables	Social Biology	Ecological fallacy. Unclear addresses for geocoding. Covariates of age, older than 50 years and overweight



Table 3. Cont.

No.	Author (Year), Country	Objective	Study Design	Sample	Statistical Test	Findings	Factors Associated	Determinants	Limitation
4	Pakzad et al. (2016), Iran	To investigate the spatial distribution of colorectal cancer in Iran	Ecological	6210 cases of CRC in 2009	Getis-Ord-Gi* (spatial statistics)	Higher incidence of CRC among men (11.31 per thousand people) than women (10.89 per thousand people) in the northern and central provinces of Iran.	Sex	Biology	Incomplete data registration. Lack of full report Poor data classification
5	Pourhoseingholi et al. (2020), Iran	To determine the distribution of CRC risk in a map with socioeconomic risk factors adjustment	Cross-sectional	21,543 CRC cases between 2005 and 2008	Generalized Linear Model (using WINBUGS program, ArcGIS v10)	Hotspot areas of CRC cluster identified for men were in the North and Western regions (mean SIR 1.92) while the Central provinces reported higher rate for women (mean SIR 1.85). Unemployment rate and mean household income had minimal impacts on CRC cluster.	Unemployment rate (mean $\pm$ SD: 11.64% $\pm$ 3.18%), Mean household income (mean $\pm$ SD: 66.46 Rials $\pm$ 12.04 Rials)	Social	Incomplete and lack of up-to-date data on population.
6	Goungounga et al. (2016), France	To compare empirically different cluster detection methods to find spatial clusters of cancer cases	Cross-sectional	3084 CRC cases between 1998 and 2007	Moran's I, the empirical Bayes index (EBI) and Potthoff-Whittinghill test (using SpODT, SaTScan an HBSM)	The socioeconomic inequalities did not affect the spatial variations of CRC incidence	Socioeconomic disadvantage as proxy by the calculation of Townsend index of deprivation. The index considers the proportion of unemployed people in the workforce, proportion of households without car, proportion of households renting and the proportion of overcrowded households. Increase in the Townsend index indicates an increase in the deprivation level of the inhabitants.	Social	Power of spatial cluster detection methods increases with the event rate
7	Roquette et al. (2019), Portugal	To describe and discuss the geographical patterns of CRC incidence and mortality in Portugal municipalities	cross sectional	37,543 CRC cases between 2007 and 2011	Global Moran's Index and Local Moran's Index (LISA), geographically weighted regression (GWR) (using ArcGIS)	The CRC incidence was relatively higher in the Norte region to women than men. Meanwhile, areas in the coastal Centro, the LVT and the Alentejo region showed markedly higher CRC cases among men than to women.	Sex	Biology	Limited data availability on other risk factors

Table 3. Cont.

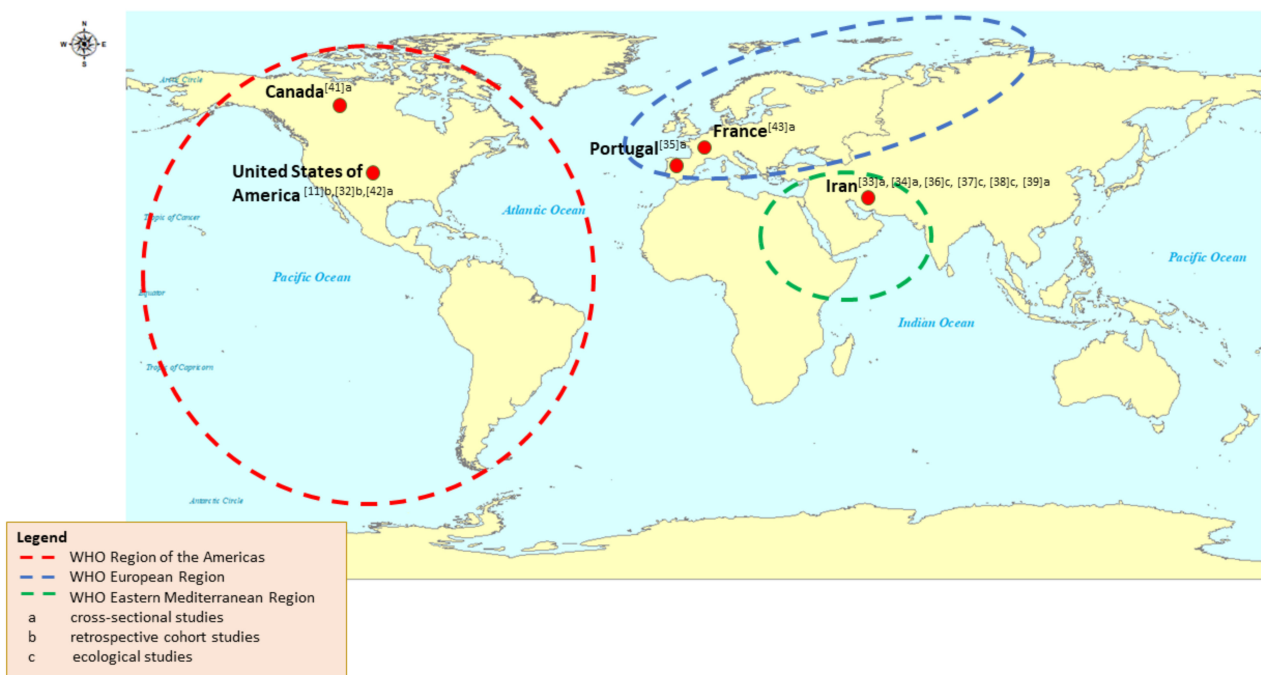
No.	Author (Year), Country	Objective	Study Design	Sample	Statistical Test	Findings	Factors Associated	Determinants	Limitation
8	Torres et al. (2018), United States	To evaluate the geographic distributions of colorectal cancer incidence among female residents in Baltimore City, Maryland and the neighborhood characteristics associated with those distributions	retrospective cohort	1120 female CRC patients between 2000 and 2010	Spatial clusters identified using Getis-Ord-Gi* statistic and local Moran's I, global ordinary least square regression model (Using STATA, ArcGIS)	Cluster spot for CRC was identified in two out of 55 Community Statistical Area (CSA) studied. The findings noted that every one percent increase in African-American residents resulted in CRC incidence increasing by 0.031 times per 1000 female residents. The CRC cluster spot experienced less crime with majority residents between the ages of 50 and 74 years old.	The 2012 Baltimore Neighborhood Indicators Alliance Vital Signs report was referred for the neighborhood characteristics. The indicators include females aged 50 to 74 years, percentage of African-American, female-headed households, percentage of households earning less than \$25,000, percentage of vacant residential properties, housing violations, number of crime incidents per 1000 residents, count of emergency call for domestic violence, teen birth, employment rates, dirty streets, tree coverage and neighborhood associations.	Social Ecology	Social determinants limited by the residential environment. No information on the length of residency on their addresses.
9	Halimi et al. (2019), Iran	To explore the spatial pattern of CRC incidence in Hamadan province, Iran	Cross-sectional	805 CRC cases during 2007 to 2014	Local Moran's I (MS Excel, Arc GIS 10.5)	High-high clusters of CRC incidence identified in Mohajeran and Lalejin areas. Majority of the CRC incidence were among men (54%) and those in the age group between 65 and 85 years.	Sex, Age	Biology	Lack of accuracy for registries. Census population data is used
10	Singh et al. (2017), Canada	To determine the variation of CRC incidence by average household income in area of residence	Cross-sectional	19,484 CRC cases between 1985 and 2012	Bayesian Poisson regression models (Using WinBUGS software, ArcGIS v10.3)	There were few small geographic areas in the southwest rural Manitoba with persistent high CRC incidence	Sex, Age, Mean annual household income, proportion of recent immigrants (since 1961 to 2001), rate of visible minority status and unemployment status.	Biology Social	Ecological analysis, hence result should be interpreted in the context of area of residence
11	Kuo et al. (2019), United States	To address spatial autocorrelation between CRC and county-level determinants	Cross sectional	2003 to 2013	Local Indicators of Spatial Association (LISA), Moran's I (Using ArcMap, SAS, GeoDaSpace software)	The location of high-high clusters identified in the north-eastern counties	Age, Adults with BMI $\geq 30$ kg/m <sup>2</sup> (Obesity), Current smoker adults who smoked at least 100 cigarettes in lifetime, Socioeconomic Status deprivation composed of education level, employment rate, income level, family and social support. Ethnicity studied include non-Hispanic Black, Hispanic, Native American and Asian. Percentage of population aged less than 65 years without health insurance. Urbanicity classified into urban, large town and rural counties.	Biology Social Ecology	Result may not generalize to other geographic areas. Lack of person-level and tumor level data

Table 3. Cont.

No.	Author (Year), Country	Objective	Study Design	Sample	Statistical Test	Findings	Factors Associated	Determinants	Limitation
12	Goshayeshi et al. (2019), Iran	To identify potential spatial factors contributing to its geographical distribution	Cross sectional	1089 CRC cases between 2016 and 2017	Local Moran's I, Ordinary Least Square Regression (Using MS Excel, ArcGIS v10.6)	CRC clusters identified in Rezashahr, Sarafrazaan and Nofel-Loshato areas. The neighborhood of CRC hotspots areas was associated with high proportion of population with 50 years and above, obesity (Body Mass Index $\geq 30$ kg/m <sup>2</sup> ), daily fibre intake ( $\leq 25$ g).	Age, Body Mass Index (BMI), daily consumption of red meat (gram), daily consumption of fibre (gram)	Biology	Study did not consider processed meats. Patients who had shifted were not included. Inability to geocode Persian addresses affect the accuracy

### 3. Results

A total of 12 studies were included in this systematic review (Table 3). Descriptive summary of included studies concerning study location, setting and study design are shown in Figure 2. All eligible studies were conducted spanning five countries that include Canada [41], United States [11,32,42], France [43], Portugal [35], and Iran [33,34,36–39]. Comparing study locations based on the WHO regions, four studies [11,32,41,42] accounted for the Region of the Americas (AMR), two studies [35,43] from the European Region (EUR) and six studies [33,34,36–39] were performed in the Eastern Mediterranean Region (EMR). The analyzed articles were published in the year 2016 to 2021. Seven of these quantitative studies resumed cross sectional study design [33–35,39,41–43], two studies were conducted using retrospective cohort [11,32], and three articles of ecological studies [36–38].



**Figure 2.** Distribution of included articles based on countries, WHO regions and study design.

The period of data collection for a quarter of the included studies dated back before 2000 [32,41,43] whereas nine more ranges between 2000 to 2017 [11,33–39,42]. The duration of retrospective data collection can be categorized into less than 10 years [11,33–39,42,43], more than 10 years [32] and more than 20 years [41]. Spatial analysis was applied to address a range of objectives (Table 3), with the most common include description of the CRC distribution, analysis of associated factors [11,32,33,36,38,39,41,42], and comparison of different cluster detection method [43].

#### 3.1. Type of Spatial Analysis

The review showed multiple types of spatial cluster statistics being used across the included studies. Seven studies [11,34,35,38,39,42,43] utilized Moran's Index to summarize the spatial autocorrelation over study area, three studies [32,36,41] used Poisson Regression Model, three studies [11,32,37] used Getis–Ord  $G_i^*$ , two studies [35,42] used local indicators of spatial association (LISA) and each one study analyzed their data using at least Besag–York–Mollie (BYM) model [38] and Generalized Linear Models [33], respectively or in combination with the others. Generally, the type of test can be classified into global, local, and focused tests according to the study hypotheses. The global cluster statistics, such as Moran's  $I$  often used to inform the existence of spatial structure of an area, not considering the point of location or the difference between different cluster [44]. Meanwhile, local

statistics such as LISA and Ord–Getis Gi, explained on the nature of the spatial dependency of a given locality and focused test (e.g., Poisson Regression Model) explore the possible clusters near potential risk factors [44].

### 3.2. Factors Associated with CRC Cluster

Evidence of clustering were abundance as most of the studies reported presence of CRC cluster in the study population. However, the outcomes were more meaningful when the studies incorporate other factors to further understand the association with CRC cluster. The review found multiple factors frequently studied such as age, sex, ethnicity, overweight and obesity, smoking, daily consumption of fruit and milk, socioeconomic status as represented by the median income level, employment status, health costs, housing violations or domestic violence, health care coverage, urbanicity, dirty streets and tree coverage. Collectively, these can be summarized into social, biology and ecology determinants (Table 4).

**Table 4.** Factors analyzed in each of the included studies.

CRC Determinants	Studies
Ecology	
Health care coverage	Kuo et al. 2019
Urbanicity	Kuo et al. 2019
Dirty streets	Torres et al. 2018
Tree coverage	Torres et al. 2018
Biology	
Age	Roquette et al. 2019; Halimi et al. 2019; Singh et al. 2017; Kuo et al. 2019; Goshayeshi et al. 2019
Sex	Pakzad et al. 2016; Roquette et al. 2019; Halimi et al. 2019; Singh et al. 2017
Ethnicity	Liu et al. 2016; Kuo et al. 2019
Overweight and Obesity	Mansori et al. 2018; Mansori et al. 2019; Kuo et al. 2019; Goshayeshi et al. 2019
Daily consumption of fruit	Mansori et al. 2018; Goshayeshi et al. 2019
Daily consumption of milk	Mansori et al. 2018; Mansori et al. 2019
Social	
Smoking	Mansori et al. 2018; Mansori et al. 2019; Kuo et al. 2019
Median household income level	Liu et al. 2016; Pourhoseingholi et al. 2020; Goungounga et al. 2016; Torres et al. 2018; Singh et al. 2017
Health costs	Mansori et al. 2018; Mansori et al. 2019
Employment status	Mansori et al. 2019; Pourhoseingholi et al. 2020; Goungounga et al. 2016; Torres et al. 2018; Kuo et al. 2019
Housing violations, domestic violence	Torres et al. 2018

When comparing the factors associated with CRC cluster in all the included studies, eight studies [11,32,33,36,38,41–43] defined the relationship of CRC cluster with social factors, another nine studies [32,34–39,41,42] explain on the biology factors, while two studies [11,42] analyzed on the ecological factor.

While CRC has been commonly linked to westernized diet and physical inactivity, less focus was given to explore on the ecological factor leveraging towards colorectal carcinogenesis. Two studies suggest that the surrounding physical environment has temporally shaped the progression of CRC cluster within an area [11,42]. High accessibility to health-care facilities was found correlated with substantial CRC screening rate, hence increase in CRC incidence and cluster [42]. Similarly, the CRC clusters were found dependent with the urbanicity level of an area [42]. The definition of an urban area by Kuo et al. (2019) was determined based on the population density of each county, which can be misleading when applied to contextual geographic variation due to ecological factor [42]. Fast food

outlets offering high-dense fat diet and processed meat were more common in urban cities, thus worth explore to predict future CRC cluster. Even though factors like dirty streets and tree coverage as proxy to the physical environments were incorporated into the scoring of the community statistical area characteristics, the role of ecology per se was not highlighted [11]. Therefore, future studies to explore on the influence of ecological factors and CRC cluster is recommended.

Eight studies [11,32,33,36,38,41–43] applied the spatial autocorrelation approach to explore the influence of social factors on CRC cluster and found out that neighborhoods with higher median household income level ranging between USD 38,040 and USD 80, 876 annually in 2011, were associated with decreased risk of both early and late-stage CRC [32]. On the other hand, in the middle-income countries where universal health coverage is an issue, the high CRC cluster pattern was observed with more frequent utilization of health services as measured through the health cost. Indirectly, the association explained the impact of socioeconomic inequalities against CRC incidence over time [36,43]. However, further information regarding the tumor stage following early detection and treatment deem important is lacking to complement the circumstances. Besides that, the aggregation analysis makes it difficult to elicit causal effect linkage on individual basis with regards to respective economic background.

The biological factors frequently analyzed in CRC spatial cluster studies were age and sex. Geographical areas with ageing population tend to form CRC cluster as compared to the younger age group [34,39,41]. However, information on the length of residency and migration activities were lacking to verify the plausible relationship between age and CRC cluster in the context of residential areas [11,42]. On the other hand, Pakzad et al. (2016) and Roquette et al. (2019) revealed specific spatial pattern of CRC cluster for men and women respectively [35,37]. The findings suggest for potential sex-specific determinants susceptible to CRC in certain areas despite the exposure to several other risk factors such as high fat diet, physical inactivity, and smoking. In areas with heterogenous ethnic population, Liu et al. (2016) reported higher CRC incidence among the Hispanics than non-Hispanics Whites and Blacks [32]. The selective trend against particular ethnicities likely supports the notion of gene–environment interaction in the progression of CRC that may arise from culturally specific dietary pattern and lifestyle. However, other factors such as the length of residency and social reciprocity should be critically considered and controlled with the native population.

Based on the findings, there is compelling evidence for future research on the interaction of ecological, biological, and social factors collectively with the geographic distribution of CRC incidence, to create area-level tailored cancer care services. Through the baseline information on the local patterns of CRC distribution, allocation of resources could be made available and planning of more targeted community intervention.

#### 4. Discussion

The review systematically identified frequent reference made to coin the population-based CRC cluster through methodological approach. Differences in methodology and statistical methods used were described to gauge better understanding of spatial cluster definition [11,45]. To date, there is little consensus on the definition of cluster pertaining to non-communicable diseases specifically colorectal cancer in the community [46]. Whereas the CDC defines a cancer cluster as “a greater than expected number of cancer cases that occurs within a group of people in a geographic area over a defined period of time”, there is notoriously vague and grey area on the definitive baseline figure for such cancer in an area [25,47]. The challenges elicited upon investigation of cancer clusters at fields had called for multiple arguments on the validity of the statistical analyses [27,48]. To overcome the potential inflated probability, this has called for alternative approach of using a standardized incidence ratio (SIR) analysis based on causal inference framework rooted for cancer cluster [49]. The statistical constructs suggested for exposure hypotheses as compared to the traditional observed cancer outcomes shed some light to more reliable findings.

Spatial cluster analysis plays an important role in quantifying geographic variation patterns [26,42,50]. It is commonly used in disease surveillance, spatial epidemiology, population genetics, landscape ecology, and many other fields, but the underlying principles are the same [51,52]. Spatial patterns are of interest to be used in cancer research to explain the link between exposure to the surrounding environments and development of cancer over more than ten years, of which the existing environment and ecosystem might have undergone drastic changes. Several approaches to the geographic pattern recognition include visualization techniques based on “eye-balling”, kernel-based methods that accentuate differences on a surface, artificial intelligence approaches and exploratory spatial data analysis (ESDA) which rely on statistical test [44].

Despite the progress in spatial statistics utilized in CRC research, the review highlighted that many studies are complacent with the traditional Moran’s I index to examine the spatial independence. The frequent usage may be related to the universal understanding of similar interpretation behavior relatively when compared to correlation coefficient [53]. With the growing interest of spatial statistics methodologies employed in various contexts, the heterogeneous geospatial studies of CRC incidence have led to difficulty in the comparison of study outcomes [32,37,38]. Spatial analysis was applied to improve understanding of a range CRC-related issues, including the distribution and determinants, the mechanisms driving the local CRC epidemiology, the effect of preventive strategies and the barriers to seek for treatment. Often, the geospatial methods have been combined with environmental factors exposure to understand the drivers of local cancer epidemiology; however, such studies remain limited for CRC in high-incidence areas [11,42,54,55].

Whereas the ecological determinants showed great significance when applied in spatial research, this factor has been lacked studied in relation to CRC. Factors such as health care coverage, urbanicity, dirty streets, and tree coverage were analyzed as the ecological determinants in the review. High accessibility of healthcare facilities offers better services including screening, thus concomitant with high CRC cluster [42,56]. Many studies linked urbanicity with higher CRC incidence partly due to the availability of health infrastructure and advance treatment options, besides the highly dense population. Walkability areas and greenness of streets have been associated with the amount and duration of physical activity [57], one of the well-known protective factors for CRC.

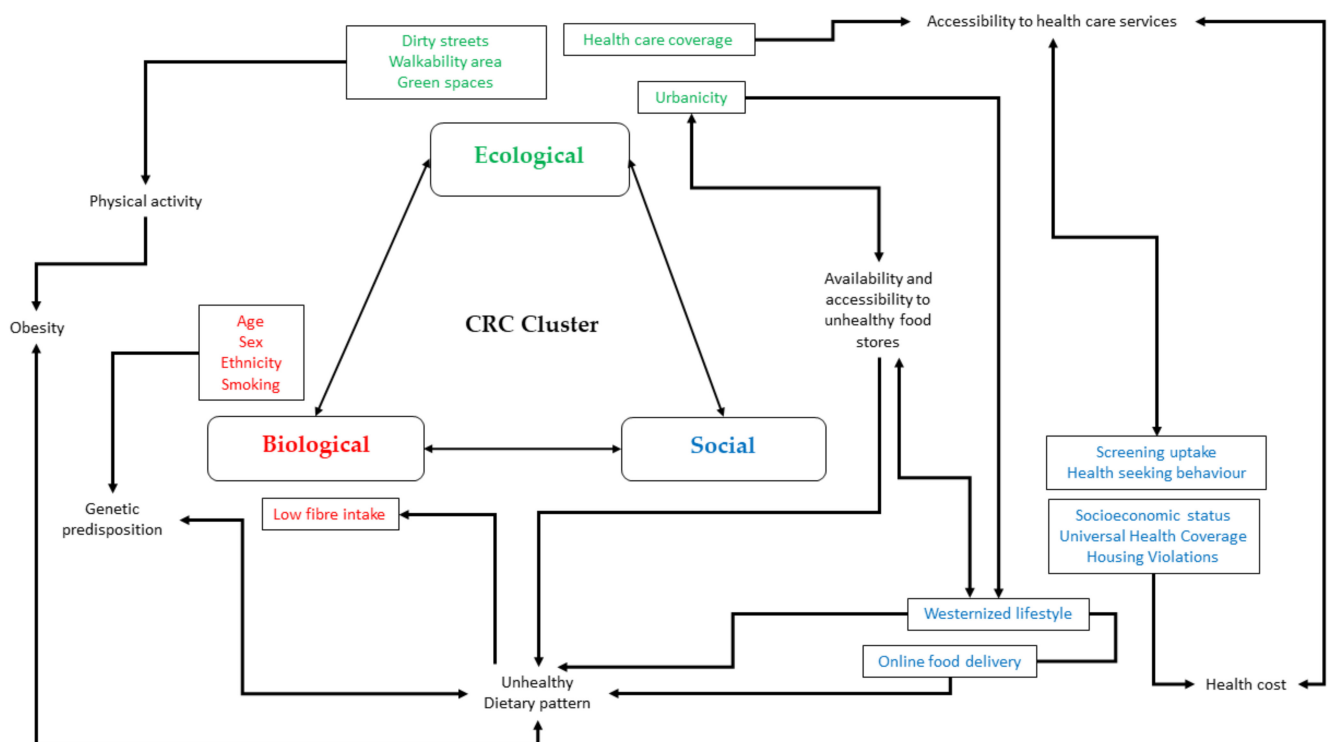
Exposure to unhealthy food environments such as the availability and accessibility to unhealthy food stores may encourage the surrounding community to have less healthy diets [54,58,59]. Likewise, the absence of green spaces or recreational parks nearby for physical activity may lead to continuous physical inactivity [19,60,61]. These are the examples of physical environment potentially instill the unhealthy lifestyle to localized settings, which may pose higher risk towards CRC in long-term. The review identified minimal studies that explore the association of the physical environment with population-based CRC cluster [11,42]. Physical environment plays important role that influence the formation of obesogenic environment, shaping the behavior and lifestyle of the population [62]. It contributes both direct and indirect pathways towards occurrence of CRC. Previous literatures highlighted the significance of ecological factors’ exposure with CRC clusters to justify public health actions and policy in the context of preventive strategies [14,54,60,61,63]. Inability to identify potential modifiable factors within the physical environment poses salient challenges towards future CRC prevention and control.

Previous literatures examined molecular genetics found associated with CRC. This includes the APC gene, K-ras family, p53, DCC, and several mismatched repair genes leading to mutations throughout the genome of affected cells. While more than 50% of sporadic CRC cases were linked to some degree of genetic mutation, the occurrence of COX-2 genetic polymorphism is particularly high among Caucasians compared to the Asians [64]. Males are more likely than females to be diagnosed as CRC across all age groups, demonstrating the role of sex in carcinogenesis [64,65].

Other social aspects such as adherence to the existing CRC screening program may provide insight to explain the existence of a local cluster pattern [66,67]. A high-incidence

area for CRC can be due to large uptake of CRC screening by the people, indicating good health awareness [68]. The behavior and attitude towards health highly influenced by the socioeconomic status [69]. Many studies supported better health outcome in countries providing universal health coverage [41]. Similarly, few recent spatial studies showed high–high clusters of CRC concentrated in urban areas compared to rural areas [11,34,42]. This can be further explained by the urbanized lifestyle that leads to more readily accessible and available online food delivery at present [70], promoting physical inactivity across all age groups in the population. Thus, future studies that examine the influence of urbanized lifestyle with the formation of CRC spatial cluster is recommended.

The ecological, biological, and social determinants have significant impacts to formation of geographical aggregated pattern of CRC incidence to an extent, when studied independently. In circumstances of the true population, most of the factors present simultaneously and possibly interact with each other, producing greater effect to increased risk of CRC. Although major studies highlighted the synergistic effect at the individual level, through animal studies supporting the gene environment interaction [71,72], there have been limited studies that examine the interaction between these factors to benefit the preventive strategies at the population level. The combined effect analysis is crucial to inform the multisectoral stakeholders on the challenging CRC burden as a shared public health issue. Figure 3 summarizes the interaction occurring between ecological, biological, and social determinants that potentially influence the existence of CRC cluster within population.



**Figure 3.** Conceptualization of interaction between ecological, biological, and social determinants influence formation of CRC cluster.

#### 4.1. Strength

The review identified potential future research area on the association of ecological, biological, and social factors with the clustering pattern of CRC incidence. With the existing knowledge on population CRC cluster driven by various sociodemographic circumstances, future studies can be designed to explore the physical and built environment across various geographic settings. Furthermore, it provides insight to the multilevel stakeholders on more specific intervention and preventive strategies tailored to the high-risk areas.



#### 4.2. Limitation

While ecological, biology, and social determinants are interrelated to cause colorectal cancer both at the individual level and community level, it is difficult to distinguish the relations of each factor independently. Spatial analysis studies focusing on cancer incidence secondary to occupation-related were not included in the review, thus limit the discussion on ecological and social influence towards cluster distribution.

Most of the reviewed articles were from middle-income settings, which may either reflect publication bias or a focus of research efforts on such settings. In high-incidence countries of the Asian region, studies with limited use of spatial analysis methods could reflect a lack of access to information resources or insufficient expertise in these settings. Nonetheless, the review revealed areas with high CRC incidence stand to gain the most from understanding of CRC spatial patterns in which clustering may be important epidemiologically.

Nearly all the models have shown significant associations between CRC cluster and demographic, socioeconomic, and risk-factor variables, although it is difficult to rule out publication bias favoring studies with positive findings. However, associations observed between CRC cluster and different factors such as sex, household income, and obesity at the population level vary across studies. These were recognized as important individual-level risk factors, highlighting the potential for ecological fallacy.

#### 5. Conclusions

In conclusion, the review identified robust evidence of CRC cluster across different geographical settings. However, attempts to examine the association of area-level determinants and CRC cluster are lacking in ecology as compared to the common biology and social attributes. Therefore, future spatial studies that incorporate physical environment (ecological) factors in this research field are warranted as guides for policymakers to plan more targeted preventive and control actions. Studies relating more than one determinant with the CRC cluster displayed potential degree of interaction, which is understudied. Future interaction analysis that incorporates the combination of ecology, biology, and social attributes may benefit to explain the trend of CRC cluster in detail, thus validating the cancer control continuum planning.

**Author Contributions:** Conceptualization, S.S.S.S. and A.M.N.; Methodology, S.S.S.S. and A.M.N.; Validation, A.M.N. and R.H.; Writing—Original Draft Preparation, S.S.S.S.; Writing—Review & Editing, A.M.N., R.H., H.-K.C. and M.R.A.H.; Visualization, A.M.N. and R.H.; Supervision, A.M.N., R.H. and M.R.A.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Ministry of Higher Education, Malaysia through Fundamental Research Grant Scheme-(FRGS/1/2021/SKK05/UKM/02/1) and The Universiti Kebangsaan Malaysia (UKM), grant number (FF-2021-121). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank The Ministry of Higher Education Malaysia and The Universiti Kebangsaan Malaysia, (UKM) for funding this study under the Fundamental Research Grant Scheme-(FRGS/1/2021/SKK05/UKM/02/1) and (FF-2021-121). We also thank the team for their commitment and tireless efforts in ensuring that manuscript was well executed.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. James, S.L.; Lucchesi, L.R.; Bisignano, C.; Castle, C.D.; Dingels, Z.V.; Fox, J.T.; Hamilton, E.B.; Liu, Z.; McCracken, D.; Nixon, M.R.; et al. Morbidity and mortality from road injuries: Results from the Global Burden of Disease Study 2017. *Inj. Prev.* **2019**, *26*, i46–i56. [CrossRef]
2. Siegel, R.L.; Torre, L.A.; Soerjomataram, I.; Hayes, R.B.; Bray, F.; Weber, T.K.; Jemal, A. Global patterns and trends in colorectal cancer incidence in young adults. *Gut* **2019**, *68*, 2179–2185. [CrossRef]
3. Rawla, P.; Sunkara, T.; Barsouk, A. Epidemiology of colorectal cancer: Incidence, mortality, survival, and risk factors. *Prz. Gastroenterol.* **2019**, *14*, 89–103. [CrossRef]
4. Ibrahim, N.R.W.; Chan, H.K.; Soelar, S.A.; Azmi, A.N.; Said, R.M.; Hassan, M.R.A. Incidence, clinico-demographic profiles and survival rates of colorectal cancer in Northern Malaysia: Comparing patients above and below 50 years of age. *Asian Pac. J. Cancer Prev.* **2020**, *21*, 1057–1061. [CrossRef] [PubMed]
5. Rogers, C.R.; Moore, J.X.; Qeadan, F.; Gu, L.Y.; Huntington, M.S.; Holowatyj, A.N. Examining factors underlying geographic disparities in early-onset colorectal cancer survival among men in the United States. *Am. J. Cancer Res.* **2020**, *10*, 1592–1607. [PubMed]
6. Aungkulanon, S.; Tangcharoensathien, V.; Shibuya, K.; Bundhamcharoen, K.; Chongsuvivatwong, V. Area-level socioeconomic deprivation and mortality differentials in Thailand: Results from principal component analysis and cluster analysis. *Int. J. Equity Health* **2017**, *16*, 117. [CrossRef]
7. Chazelas, E.; Srour, B.; Desmetz, E.; Kesse-Guyot, E.; Julia, C.; Deschamps, V.; Druesne-Pecollo, N.; Galan, P.; Hercberg, S.; Latino-Martel, P.; et al. Sugary drink consumption and risk of cancer: Results from NutriNet-Santé prospective cohort. *BMJ* **2019**, *366*, l2408. [CrossRef] [PubMed]
8. Domingo, J.L.; Nadal, M. Carcinogenicity of consumption of red meat and processed meat: A review of scientific news since the IARC decision. *Food Chem. Toxicol.* **2017**, *105*, 256–261. [CrossRef] [PubMed]
9. Tyrrell, J.; Wood, A.R.; Ames, R.M.; Yaghootkar, H.; Beaumont, R.N.; Jones, S.E.; Tuke, M.A.; Ruth, K.S.; Freathy, R.M.; Smith, G.D.; et al. Gene-obesogenic environment interactions in the UK Biobank study. *Int. J. Epidemiol.* **2017**, *46*, 559–575. [CrossRef]
10. Murphy, N.; Moreno, V.; Hughes, D.V.; Vodicka, L.; Vodicka, P.; Aglago, E.K.; Gunter, M.J.; Jenab, M. Lifestyle and dietary environmental factors in colorectal cancer susceptibility. *Mol. Aspects Med.* **2019**, *69*, 2–9. [CrossRef]
11. Torres, A.Z.; Phelan-Emrick, D.; Castillo-Salgado, C. Evaluating Neighborhood Correlates and Geospatial Distribution of Breast, Cervical, and Colorectal Cancer Incidence. *Front. Oncol.* **2018**, *8*, 1–10. [CrossRef] [PubMed]
12. Andrilla, C.H.A.; Moore, T.E.; Man Wong, K.; Evans, D.V. Investigating the Impact of Geographic Location on Colorectal Cancer Stage at Diagnosis: A National Study of the SEER Cancer Registry. *J. Rural Health* **2020**, *36*, 316–325. [CrossRef] [PubMed]
13. Wong, M.C.; Ding, H.; Wang, J.; Chan, P.S.; Huang, J. Prevalence and risk factors of colorectal cancer in Asia. *Intest. Res.* **2019**, *17*, 317–329. [CrossRef] [PubMed]
14. Shvetsov, Y.B.; Shariff-Marco, S.; Yang, J.; Conroy, S.M.; Canchola, A.J.; Albright, C.L.; Park, S.; Monroe, K.R.; Marchand, L.L.; Gomez, S.L.; et al. Association of change in the neighborhood obesogenic environment with colorectal cancer risk: The Multiethnic Cohort Study. *SSM—Popul. Health* **2020**, *10*, 100532. [CrossRef]
15. Keum, N.N.; Giovannucci, E. Global burden of colorectal cancer: Emerging trends, risk factors and prevention strategies. *Nat. Rev. Gastroenterol. Hepatol.* **2019**, *16*, 713–732. [CrossRef]
16. Dekker, E.; Tanis, P.J.; Vleugels, J.L.A.; Kasi, P.M.; Wallace, M.B. Colorectal cancer. *Lancet* **2019**, *394*, 1467–1480. [CrossRef]
17. Stracci, F.; Gili, A.; Naldini, G.; Gianfredi, V.; Malaspina, M.; Passamonti, B.; Bianconi, F. Geospatial analysis of the influence of family doctor on colorectal cancer screening adherence. *PLoS ONE* **2019**, *14*, e0222396. [CrossRef]
18. Glover, M.; Mansoor, E.; Panhwar, M.; Parasa, S.; Cooper, G.S. Epidemiology of Colorectal Cancer in Average Risk Adults 20–39 Years of Age: A Population-Based National Study. *Dig. Dis. Sci.* **2019**, *64*, 3602–3609. [CrossRef]
19. Ling, O.H.L.; Mohamed Musthafa, S.N.A.; Hamzah, M.S.; Marzukhi, M.A.; Mabahwi, N.A. Health and Physical Activity in Urban Neighbourhoods. Case Study: Shah Alam City, Selangor, Malaysia. *Built Environ. J.* **2020**, *17*, 11. [CrossRef]
20. Mahendra, A.; Polsky, J.Y.; Robitaille, É.; Lefebvre, M.; McBrien, T.; Minaker, L.M. Geographic retail food environment measures for use in public health. *Health Promot. Chronic Dis. Prev. Can.* **2017**, *37*, 357–362. [CrossRef]
21. Quintero, J.; Pulido, N.R.; Logan, J.; Ant, T.; Bruce, J.; Carrasquilla, G. Effectiveness of an intervention for *Aedes aegypti* control scaled-up under an intersectoral approach in a Colombian city hyperendemic for dengue virus. *PLoS ONE* **2020**, *15*, e0230486. [CrossRef]
22. Ulijaszek, S.; McLennan, A.; Graff, H. Interactions: Lessons from Obesity. In *A Companion to the Anthropology of Environmental Health*; John Wiley & Sons: Hoboken, NJ, USA, 2016; p. 85.
23. Rushton, G.; Peleg, I.; Banerjee, A.; Smith, G.; West, M. Analyzing geographic patterns of disease incidence: Rates of late-stage colorectal cancer in Iowa. *J. Med. Syst.* **2004**, *28*, 223–236. [CrossRef]
24. Geoffrey, M.J.; Dunrie, A.G. Local clustering in breast, lung and colorectal cancer in Long Island, New York. *Int. J. Health Geogr.* **2003**, *2*, 1–12.
25. National Cancer Institute: Cancer Clusters. 2018. Available online: <https://www.cancer.gov/about-cancer/causes-prevention/risk/substances/cancer-clusters-fact-sheet> (accessed on 20 July 2021).
26. He, R.; Zhu, B.; Liu, J.; Zhang, N.; Zhang, W.H.; Mao, Y. Women’s cancers in China: A spatio-temporal epidemiology analysis. *BMC Women’s Health* **2021**, *21*, 116. [CrossRef]

27. Cho, J.; You, S.C.; Lee, S.; Park, D.; Park, B.; Hripcsak, G.; Park, R.W. Application of epidemiological geographic information system: An open-source spatial analysis tool based on the OMOP common data model. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7824. [[CrossRef](#)]
28. Page, M.; McKenzie, J.; Bossuyt, P.; Boutron, I.; Hoffmann, T.; Mulrow, C. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [[CrossRef](#)]
29. Downes, M.J.; Brennan, M.L.; Williams, H.C.; Dean, R.S. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open* **2016**, *6*, 1–7. [[CrossRef](#)] [[PubMed](#)]
30. Lockwood, C.; Munn, Z.; Porritt, K. Qualitative research synthesis: Methodological guidance for systematic reviewers utilizing meta-aggregation. *JBI Evid. Implement.* **2015**, *13*, 179–187. [[CrossRef](#)] [[PubMed](#)]
31. Higgins, J.P.; Thomas, J.; Chandler, J.; Cumpston, M.; Li, T.; Page, M.J.; Welch, V.A. *Cochrane Handbook for Systematic Reviews of Interventions*; John Wiley & Sons: Hoboken, NJ, USA, 2019.
32. Liu, Z.; Zhang, K.; Du, X.L. Risks of developing breast and colorectal cancer in association with incomes and geographic locations in Texas: A retrospective cohort study. *BMC Cancer* **2016**, *16*, 294. [[CrossRef](#)]
33. Pourhoseingholi, M.A.; Nahafimehr, H.; Kavousi, A.; Pasharavesh, L.; Khanabadi, B. The spatial distribution of colorectal cancer relative risk in Iran: A nationwide spatial study. *Gastroenterol. Hepatol. Bed Bench* **2020**, *13*, S40–S46. [[PubMed](#)]
34. Halimi, L.; Bagheri, N.; Hoseini, B.; Hashtarkhani, S. Spatial Analysis of Colorectal Cancer Incidence in Hamadan Province, Iran: A Retrospective Cross-Sectional Study. *Appl. Spat. Anal. Policy* **2020**, *13*, 293–303. [[CrossRef](#)]
35. Roquette, R.; Painho, M.; Nunes, B. Geographical patterns of the incidence and mortality of colorectal cancer in mainland Portugal municipalities (2007–2011). *BMC Cancer* **2019**, *19*, 512. [[CrossRef](#)]
36. Mansori, K.; Solaymani-Dodaran, M.; Mosavi-Jarrahi, A.; Motlagh, A.G.; Salehi, M.; Delavari, A.; Hosseini, A.; Asadi-Lari, M. Determination of effective factors on geographic distribution of the incidence of colorectal cancer in Tehran using geographically weighted Poisson regression model. *Med. J. Islam. Repub. Iran* **2019**, *33*, 23. [[CrossRef](#)] [[PubMed](#)]
37. Pakzad, R.; Moudi, A.; Pournamdar, Z.; Pakzad, I.; Mohammadian-Hafshejani, A.; Momenimovahed, Z.; Salehiniya, H.; Towhidi, F.; Makhsosi, B.R. Spatial analysis of colorectal cancer in Iran. *Asian Pac. J. Cancer Prev.* **2016**, *17*, 53–57. [[CrossRef](#)] [[PubMed](#)]
38. Mansori, K.; Solaymani-Dodaran, M.; Mosavi-Jarrah, A.; Motlagh, A.G.; Salehi, M.; Delavari, A.; Asadi-Lari, M. Spatial inequalities in the incidence of colorectal cancer and associated factors in the neighborhoods of Tehran, Iran: Bayesian spatial models. *J. Prev. Med. Public Health* **2018**, *51*, 33–40. [[CrossRef](#)]
39. Goshayeshi, L.; Purahmadi, A.; Ghayour-Mobarhan, M.; Hashtarkhani, S.; Karimian, S.; Dastjerdi, R.S.; Eghbali, B.; Seyfi, E.; Kiani, B. Colorectal cancer risk factors in North-Eastern Iran: A retrospective cross-sectional study based on geographical information systems, spatial autocorrelation and regression analysis. *Geospat. Health* **2019**, *14*, 219–228. [[CrossRef](#)] [[PubMed](#)]
40. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [[CrossRef](#)]
41. Singh, H.; Nugent, Z.; Decker, K.; Demers, A.; Samaddar, J.; Torabi, M. Geographic variation and factors associated with colorectal cancer incidence in Manitoba. *Can. J. Public Health* **2017**, *108*, e558–e564. [[CrossRef](#)]
42. Kuo, T.M.; Meyer, A.M.; Baggett, C.D.; Olshan, A.F. Examining determinants of geographic variation in colorectal cancer mortality in North Carolina: A spatial analysis approach. *Cancer Epidemiol.* **2019**, *59*, 8–14. [[CrossRef](#)]
43. Goungounga, J.A.; Gaudart, J.; Colonna, M.; Giorgi, R. Impact of socioeconomic inequalities on geographic disparities in cancer incidence: Comparison of methods for spatial disease mapping. *BMC Med. Res. Methodol.* **2016**, *16*, 136. [[CrossRef](#)]
44. Jacquez, G. Spatial Cluster Analysis. In *The Handbook of Geographic Information Science*; Blackwell Publishing: Hoboken, NJ, USA, 2008; pp. 395–416.
45. Congdon, P. Variations in obesity rates between US counties: Impacts of activity access, food environments, and settlement patterns. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1023. [[CrossRef](#)]
46. Blangiardo, M.; Bouliერი, A.; Diggle, P.; Piel, F.B.; Shaddick, G.; Elliott, P. Advances in spatiotemporal models for non-communicable disease surveillance. *Int. J. Epidemiol.* **2020**, *49*, I26–I37. [[CrossRef](#)]
47. Centers for Disease Control and Prevention. *Investigating Suspected Cancer Clusters and Responding to Community Concerns: Guidelines from CDC and the Council of State and Territorial Epidemiologists*; Morbidity and Mortality Weekly Report 62(RR08); Centers for Disease Control and Prevention: Atlanta, GA, USA, 2013; pp. 1–14.
48. Roquette, R.; Painho, M.; Nunes, B. Spatial epidemiology of cancer: A review of data sources, methods and risk factors. *Geospat. Health* **2017**, *12*, 2000–2010. [[CrossRef](#)]
49. Nethery, R.C.; Yang, Y.; Brown, A.J.; Dominici, F. A causal inference framework for cancer cluster investigations using publicly available data. *J. R. Stat. Soc. Ser. A Stat. Soc.* **2020**, *183*, 1253–1272. [[CrossRef](#)]
50. Amin, R.W.; Fritsch, B.A.; Retzliff, J.E. Spatial Clusters of Breast Cancer Mortality and Incidence in the Contiguous USA: 2000–2014. *J. Gen. Intern. Med.* **2019**, *34*, 412–419. [[CrossRef](#)]
51. Tian, Y.; Li, J.; Zhou, T.; Tong, D.; Chi, S.; Kong, X.; Ding, K.; Li, J. Spatially varying effects of predictors for the survival prediction of nonmetastatic colorectal Cancer. *BMC Cancer* **2018**, *18*, 1084. [[CrossRef](#)]
52. Amin, R.W.; Rivera, B. A spatial study of oral & pharynx cancer mortality and incidence in the U.S.A.: 2000–2015. *Sci. Total Environ.* **2020**, *713*, 136688. [[PubMed](#)]
53. Tillé, Y.; Dickson, M.M.; Espa, G.; Giuliani, D. Measuring the spatial balance of a sample: A new measure based on Moran's I index. *Spat. Stat.* **2018**, *23*, 182–192. [[CrossRef](#)]

54. Gibson, D.C.; Prochaska, J.D.; Yu, X.; Kaul, S. An examination between census tract unhealthy food availability and colorectal cancer incidence. *Cancer Epidemiol.* **2020**, *67*, 101761. [[CrossRef](#)] [[PubMed](#)]
55. Hossain, F.; Danos, D.; Prakash, O.; Gilliland, A.; Ferguson, T.F.; Simonsen, N.; Leonardi, C.; Yu, Q.; Wu, X.C.; Miele, L.; et al. Neighborhood social determinants of triple negative breast cancer. *Front. Public Health* **2019**, *7*, 1–8. [[CrossRef](#)] [[PubMed](#)]
56. Fujita, M.; Sato, Y.; Nagashima, K.; Takahashi, S.; Hata, A. Impact of geographic accessibility on utilization of the annual health check-ups by income level in Japan: A multilevel analysis. *PLoS ONE* **2017**, *12*, e0177091. [[CrossRef](#)]
57. Mueller, W.; Wilkinson, P.; Milner, J.; Vardoulakis, S.; Steinle, S.; Parkka, J.; Parmes, E.; Cluitmans, L.; Kuijpers, E.; Pronk, A.; et al. Neighbourhood and path-based greenspace in three European countries: Associations with objective physical activity. *BMC Public Health* **2021**, *21*, 282. [[CrossRef](#)] [[PubMed](#)]
58. Wende, M.E.; Stowe, E.W.; Eberth, J.M.; McLain, A.C.; Liese, A.D.; Breneman, C.B.; Josey, M.J.; Hughey, S.M.; Kaczynski, A.T. Spatial clustering patterns and regional variations for food and physical activity environments across the United States. *Int. J. Environ. Health Res.* **2020**, *00*, 1–15. [[CrossRef](#)] [[PubMed](#)]
59. Janssen, H.G.; Davies, H.G.; Richardson, L.D.; Stevenson, L. Determinants of takeaway and fast-food consumption: A narrative review. *Nutr. Res. Rev.* **2018**, *31*, 16–34. [[CrossRef](#)]
60. Wray, A.J.D.; Minaker, L.M. Is cancer prevention influenced by the built environment? A multidisciplinary scoping review. *Cancer* **2019**, *125*, 3299–3311. [[CrossRef](#)]
61. Mabry, R.M.; Al Siyabi, H.; Kannan, M.; Al Siyabi, A. Physical activity and the built environment: Perceptions of policy-makers in Oman. *Health Promot. Int.* **2020**, *35*, 762–770. [[CrossRef](#)]
62. Caperon, L.; Arjyal, A.; Puja, K.C.; Kuikel, J.; Newell, J.; Peters, R.; Prestwich, A.; King, R. Developing a socio-ecological model of dietary behaviour for people living with diabetes or high blood glucose levels in urban Nepal: A qualitative investigation. *PLoS ONE* **2019**, *14*, e0214142. [[CrossRef](#)]
63. Yuen, J.W.; Chang, K.K.; Wong, F.K.; Wong, F.Y.; Siu, J.Y.; Ho, H.C.; Wong, M.S.; Ho, J.; Chan, K.L.; Yang, L. Influence of urban green space and facility accessibility on exercise and healthy diet in Hong Kong. *Int. J. Environ. Res. Public Health* **2019**, *16*, 1514. [[CrossRef](#)]
64. Zhang, Y.C.; Zhao, H.; Chen, C.; Ali, M.A. COX-2 gene rs689466 polymorphism is associated with increased risk of colorectal cancer among Caucasians: A meta-Analysis. *World J. Surg. Oncol.* **2020**, *18*, 1–9. [[CrossRef](#)] [[PubMed](#)]
65. Gausman, V.; Dornblaser, D.; Anand, S.; Hayes, R.B.; O'Connell, K.; Du, M.; Liang, P.S. Risk Factors Associated with Early-Onset Colorectal Cancer. *Clin. Gastroenterol. Hepatol.* **2020**, *18*, 2752–2759. [[CrossRef](#)] [[PubMed](#)]
66. Zhan, F.B.; Morshed, N.; Kluz, N.; Candelaria, B.; Baykal-Caglar, E.; Khurshid, A.; Pignone, M.P. Spatial Insights for Understanding Colorectal Cancer Screening in Disproportionately Affected Populations, Central Texas, 2019. *Prev. Chronic Dis.* **2021**, *18*, 1–7. [[CrossRef](#)] [[PubMed](#)]
67. Carethers, J.M.; Doubeni, C.A. Causes of Socioeconomic Disparities in Colorectal Cancer and Intervention Framework and Strategies. *Gastroenterology* **2020**, *158*, 354–367. [[CrossRef](#)]
68. Danos, D.M.; Ferguson, T.F.; Simonsen, N.R.; Leonardi, C.; Yu, Q.; Wu, X.C.; Scribner, R.A. Neighborhood disadvantage and racial disparities in colorectal cancer incidence: A population-based study in Louisiana. *Ann. Epidemiol.* **2018**, *28*, 316–321. [[CrossRef](#)] [[PubMed](#)]
69. Ni, K.; O'Connell, K.; Anand, S.; Yakoubovitch, S.C.; Kwon, S.C.; De Latour, R.A.; Wallach, A.B.; Sherman, S.E.; Du, M.; Liang, P.S. Low colorectal cancer screening uptake and persistent disparities in an underserved urban population. *Cancer Prev. Res.* **2020**, *13*, 395–402. [[CrossRef](#)] [[PubMed](#)]
70. Maimaiti, M.; Ma, X.; Zhao, X.; Jia, M.; Li, J.; Yang, M.; Ru, Y.; Yang, F.; Wang, N.; Zhu, S. Multiplicity and complexity of food environment in China: Full-scale field census of food outlets in a typical district. *Eur. J. Clin. Nutr.* **2020**, *74*, 397–408. [[CrossRef](#)] [[PubMed](#)]
71. Pan, T.; Liu, J.; Xu, S.; Yu, Q.; Wang, H.; Sun, H.; Wu, J.; Zhu, Y.; Zhou, J.; Zhu, Y. ANKRD22, a novel tumor microenvironment-induced mitochondrial protein promotes metabolic reprogramming of colorectal cancer cells. *Theranostics* **2020**, *10*, 516–536. [[CrossRef](#)]
72. Qi, L.; Zhang, Y.; Zhang, W.; Wang, Y.; Han, Y.; Ding, Y. The inhibition of colorectal cancer growth by the natural product macrocarpal I. *Free Radic. Biol. Med.* **2021**, *162*, 383–391. [[CrossRef](#)]