IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 29 November 2023, accepted 21 January 2024, date of publication 30 January 2024, date of current version 6 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3359638

RESEARCH ARTICLE

Taste Mapping: Navigating the Spatiotemporal Link Between Diet and Colorectal Cancer

QINYAO LUO[1], YU LIU[2,3], MINYAN BI[4], XI KUAI[5], QIN TIAN[2], YUKAI SUN[6,7], AND SIDA ZHUANG[8]

[1]School of Geosciences and Info-Physics, Central South University, Changsha 410083, China
[2]Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, Shenzhen 518040, China
[3]Institute of Environment and Development, Guangdong Academy of Social Sciences, Guangzhou 510635, China
[4]SFMAP Technology (Shenzhen) Ltd., Shenzhen 518000, China
[5]Research Institute for Smart Cities, School of Architecture and Urban Planning, Shenzhen University, Shenzhen 518063, China
[6]Experimental and Clinical Research Center, Charité—University Medicine Berlin, 13125 Berlin, Germany
[7]Max-Delbrück-Center for Molecular Medicine AG Translational Oncology of Solid Tumors, 13125 Berlin, Germany
[8]Geography Department, Humboldt-Universittät zu Berlin, 10099 Berlin, Germany

Corresponding author: Yu Liu (liuyu_0201@163.com)

This work was supported in part by the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, under Grant KF-2022-07-007; in part by the Guangzhou Philosophy and Social Science Planning Project under Grant 2023GZQN60; in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515010117; in part by the Shenzhen Science and Technology Program under Grant 20220810160944001; and in part by the National Natural Science Foundation of China under Grant 41901325.

ABSTRACT Colorectal cancer is a major public health problem for its contribution to morbidity and mortality globally, and one of its major inducements is diet. Despite the achievements in the correlation studies about dietary habit and colorectal cancer, as an important component of dietary habit, taste preference has not received enough attention, its quantitative relationship with colorectal cancer has not been fully revealed. Using data from crowdsourced recipes, restaurant points of interest, and regional colorectal cancer mortality rates, this study quantitatively analysed the effects of different tastes on colorectal cancer, established a spatio-temporal link between taste and the disease, elucidating taste's role in the spatial distribution of colorectal cancer. According to the analysis, Chinese taste preference presents significant spatial heterogeneity, and Chinese cuisines most favour 'Spicy' which makes up the top two risk factors of colorectal cancer with 'Pungent'. Also, the interactions of Salty' with other tastes might increase colorectal cancer risk, the same goes for 'Oily'. The proposed quantitative analysis method based on crowdsourcing data can be applied to researches on other diseases, and the study can provide a new method and a new perspective for relevant studies.

INDEX TERMS Risk factor, taste, spatiotemporal distribution, crowdsourcing data.

I. INTRODUCTION

As one of the leading causes of morbidity and mortality worldwide, colorectal cancer (CRC) has become a major public health problem [1]. Approximately two million CRC cases are confirmed each year, with nearly one million reported deaths. Of note, middle-aged and elderly people make up more than eighty percent of the recorded cases [2], [3]. Considering the population aging, the prevalence of CRC is expected to increase significantly in the next few years. For individuals, CRC brings both serious physical injury and massive psychological damage [4], [5]. For economies and society, CRC leads to serious problems, such as productivity impairment and soaring medical expenses, costing billions of dollars annually [4], [6]. Therefore, CRC poses a tremendous personal, economic and social burden. It's worth mentioning that China is the biggest contributor to the global burden of CRC due to its large population [7]. Increasing CRC cases

The associate editor coordinating the review of this manuscript and approving it for publication was Jon Atli Benediktsson.

has been a tough challenge for China to overcome, hence, in-depth study on the characteristics of CRC in China becomes necessary [8], [9].

CRC is a chronic disease closely associated with social, biological and behavioural factors [8], [10], [11]. Eating habits has been listed as a significant contributor to CRC [4], [12]. Its strong association with CRC has been confirmed by previous studies [13], [14], [15], [16], [17], [18], [19], [20], [21], some studies also imply that the taste preferences may have influence on CRC [22]. However, most of these studies rely on questionnaires which are usually limited in sample size. More importantly, the reliance on questionnaires will inevitably bring about problems like selection bias and survival bias. Considering these issues, crowdsourcing data might be a solution for its advantages in sampling, content and balance [23], [24]. Given the above, this study chose China as the study area and utilized crowdsourcing data to carry out quantitative analyses of the taste preferences of people in different regions. The analyses are aimed at detecting a quantitative relationship between taste preference and CRC.

The significance of this study mainly lies in three aspects. First, for departments without access to the latest social survey data or clinical data from medical academies, this study may help them formulate better intervention measures. Second, the results presented in the paper can be used by relevant researches as references or supporting materials. Last, this paper presents a crowdsourcing data-driven study on the relationship between Chinese dietary preference and CRC from the perspective of demography. The study not only estimates the quantitative expression of the relationship but also utilizes crowdsourcing data to carry out a new attempt on public health information mining which has certain practical significance.

II. BACKGROUND

A. ASSOCIATION BETWEEN TASTE AND CRC

According to a report from the World Health Organization (WHO), unhealthy eating habits, long-term smoking, excessive alcohol consumption, and irregular sleep are the top four contributors to CRC. Among these, eating habits are paramount [4], [12]. Therefore, a plethora of scholars have explored the impacts of several essential nutrients, specific foods as well as overall diet structure and nutritional balance [13], [14], [15], [16]. According to their research findings, low-fat foods [14], dietary products [15], whole grains, vegetables, fruit [16] and nutrients like dietary fibre [13], calcium, vitamin D [16] are all associated with lower risk of CRC, therefore, dietary modification is potential to reduce CRC incidence [16]. In addition, the strong correlation between culinary preferences and chronic conditions has also been proven through relevant tests in recruited volunteers. For instance, *Vecchia et al.* carried out a case-control study on 953 cases of histologically confirmed colon cancer and 633 of rectal cancer in Northern Italy. In that study, relevant information on factors such as food consumption, lifestyle habits was obtained by a structured questionnaire, and the statistics show that the consumption of overly sweet food positively correlates with the multivariate relative risks of both colon and rectal cancer [25]. Besides, through questionnaire survey, such association has been also observed in other health outcomes like urinary calculi [26], diabetes [27] and etc. Moreover, preserved foods with high salt content, saturated/animal fats [17], [18], cholesterol [19], [20], high sugar foods [17], [21], spicy foods [17] and pungent spices [21]have been found by a number of studies to have a significant positive association with CRC risk. This implies a qualitative relationship between CRC and tastes, such as 'Salty', 'Oily', 'Sweet','Spicy' and 'Pungent' [22].

III. STUDY AREA

For the study area, we selected 33 of the 34 Chinese first-level administrative districts based on information from the China Urban Statistical Yearbook (Figure 1). The 33 selected districts include 22 provinces, 5 autonomous regions (Inner Mongolia, Tibet, Ningxia, Xinjiang, Guangxi), 4 municipalities directly under the central government (Beijing, Tianjin, Shanghai, Chongqing) and 2 special administrative regions (Hong Kong, Macao). Considering data accessibility, Taiwan was excluded from the study. Over 98% of the Chinese population of all ages lives in the 33 selected districts; therefore, the study results have extensive applicability [28].

IV. METHODOLOGY

As Figure 2 shows, this study was implemented through eight key steps: (1)fetching points of interest (POI) from Amap and relevant culinary information from cuisine websites with crawlers for basics, such as restaurants' locations and names, as well as detailed information about each dish, including its corresponding cuisine, recipe, and ingredients; (2) labelling the taste of each dish according to its ingredients, then perform statistics of Cuisine_Taste to produce the list of Cuisine_Taste; (3) getting the regional statistical distribution of the POI of restaurants of different cuisines, calculating the proportion of POI of each cuisine, and then obtain the list of Region_Cuisine; (4) producing the list of Region_Taste on the basis of the lists of Cuisine_Taste and Region_Cuisine; (5) visualizing and analysing the spatial distribution pattern of taste using the spatial analysis of GIS; (6) evaluating the relative importance of the influence of each different taste on CRC with LMG model; (7) conducting regression analysis of taste and interaction evaluation of impact factors utilizing Geodector; (8) analysing the results in detail.

A. LMG MODEL

Lindeman, Merenda and Gold proposed LMG measure in 1980 and named it after the English acronym of their names. LMG measure quantifies relative importance of impact factors in a given quantized regression model [29], since it takes full account of both the independent contribution and interactive contribution of all impact factors without the need for normalized correction or nonnegative correction, it is one

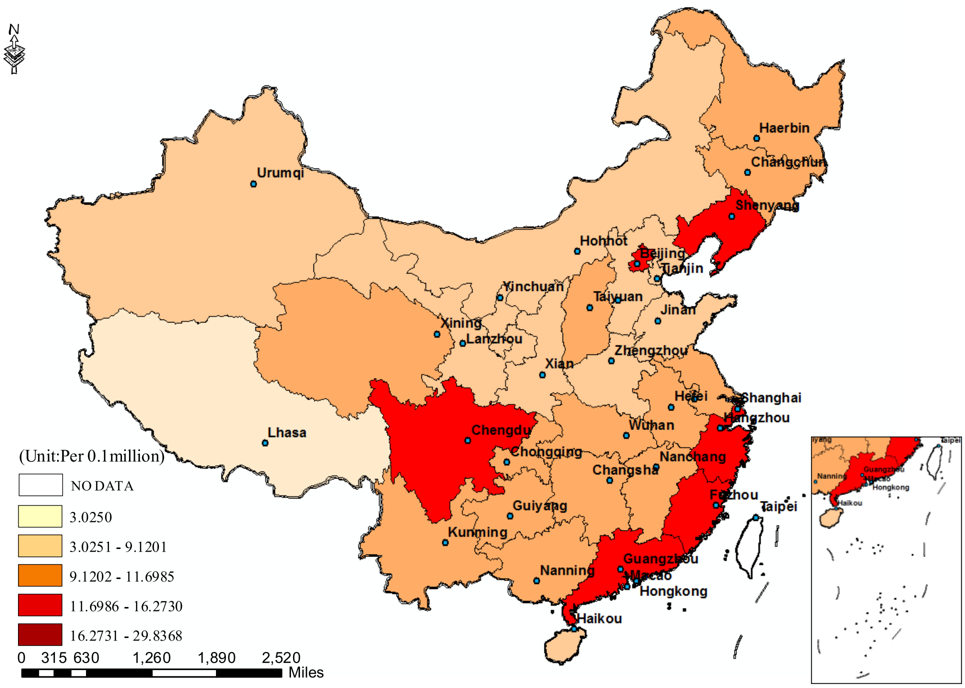

FIGURE 1. Spatial distribution of colorectal cancer mortality in china.

of the most credible measures for relative importance [30], [31]. Considering that, this study adopts the LMG model to evaluate relative importance of each taste on CRC.

LMG model considers the response variable Y, namely CRC as a linear function of n impact factors $x_1, x_2, \ldots, x_n$:

$$Y = \sum\nolimits_{i=1}^{n} \beta_i x_i + \varepsilon \tag{1}$$

where $\beta_i(i = 1, 2, \ldots, n)$ indicate the model's coefficients, and ε denotes the error term; $\beta_i(i = 1, 2, \ldots, n)$ are estimated using the collected data. With the estimations of $\beta_i(i = 1, 2, \ldots, n)$, $\delta_i\ (i = 1, 2, \ldots, n)$, the value of Y can be predicted as:

$$\hat{Y} = \sum\nolimits_{i=1}^{n} \delta_i x_i \tag{2}$$

where $\hat{Y}$ denotes the predicted value of Y.

LMG model assumes the sum of each impactor's relative importance to equal the proportion of $Var\left(\hat{Y}\right)$ on $Var\,(Y)$, which is generally represented by R^2. That can be summarized as:

$$\mathrm{Var}\left(\hat{\mathrm{Y}}\right) \Big/ \mathrm{Var}\,(\mathrm{Y}) = \sum\nolimits_{i=1}^{n} \mathrm{LMG}\,(x_i) \tag{3}$$

$$\mathrm{LMG}\,(\mathrm{x_i}) = \frac{1}{n!} \sum\nolimits_{r_{permutation}} \mathrm{SVAR}(x_i|s_i(r)) \tag{4}$$

$\mathrm{LMG}(x_i)$ refers to the relative importance of x_i, r denotes the order in which x_i enters the linear model, and its permutation is expressed as $r_{permutation} = (r_1, r_{2!}, \ldots, r_{n!})$. $s_i(r)$ denotes the set of impact factors that enter the model before x_i in the $r_{permutation}$, SVAR denotes the increase in R^2 caused by the addition of x_i when the model already contains the factors in $s_i(r)$. $\mathrm{LMG}(x_i)$ equals the average of the increase in R^2 caused by the entry of x_i into the model for all permutations [32].

B. GEODECTOR

The Geodector is a model proposed by Wang J. F. and Xu C. Ds. on the basis of the power of the determinant. This model combines GIS technology and set theory and can be used to unveil the spatial differentiation of natural phenomena and relevant driving factors. The main characteristic of the Geodector is its lack of omission, namely, it will not miss any existing relationships among given factors [33]. The model is composed of four different detectors: differentiation and factor detector, risk detector, ecological detector and interaction detector. Recently, Geodector has been applied extensively in the fields of medical science [34], [35], [36], social economics [37], [38], [39], hazard assessment [40], [41], [42], and land utilization [43], [44], [45]. Its mathematical expression is shown in Equation (5):

$$Q_{D,H} = 1 - \frac{1}{n\sigma_H^2} \sum\nolimits_{i=1}^{n} n_{D,i}\sigma_{H_{D,i}}^2 \tag{5}$$

where D refers to the impact factor associated with CRC, $Q_{D,H}$ is the indicator of the explanatory power of D, the value of $Q_{D,H}$ is between 0 and 1, $n_{D,i}$ denotes the sample number of the i^{th} parcel, and $\sigma_{H,i}^2$ is the variance in the mortality of CRC in a parcel (a parcel corresponds to a region).

CRC is assumed to be inextricably linked with taste preference in this study, which may have resulted from interactions

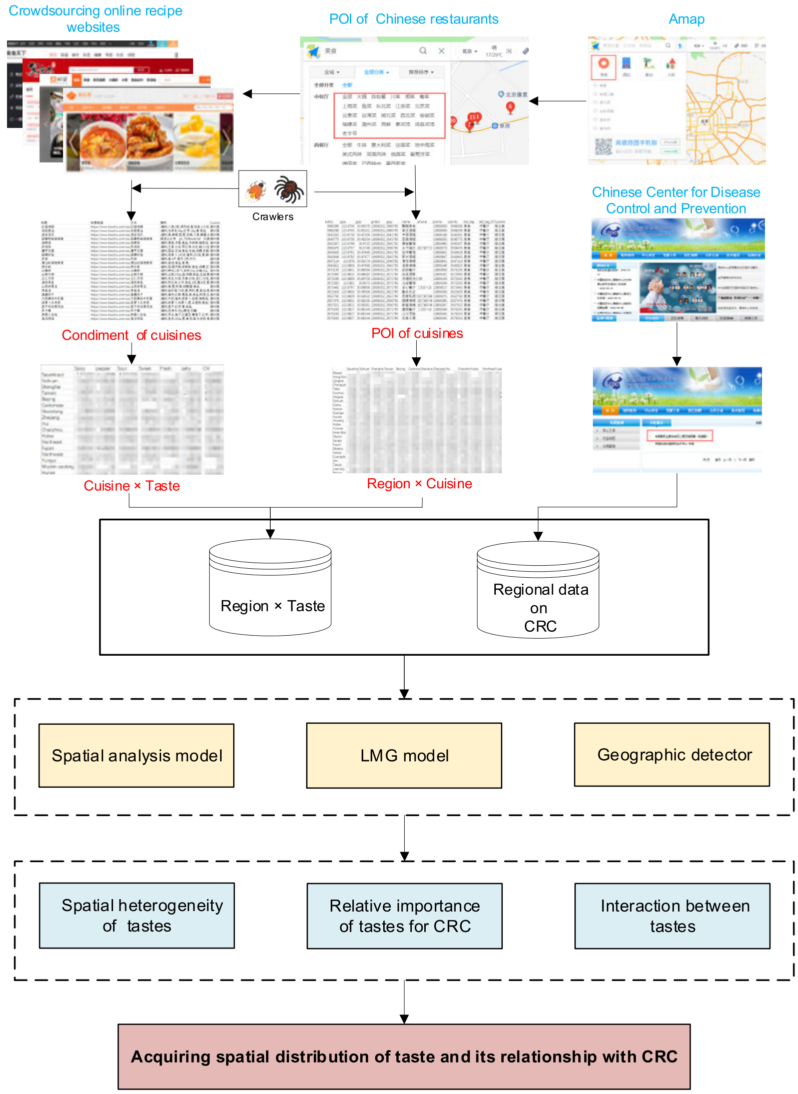

FIGURE 2. Workflow for the navigation of spatiotemporal link between taste and CRC.

TABLE 1. Descriptions and interaction relationships.

Legends	Descriptions	Interactions
	$Q(x_1 \cap x_2) < \min\,[Q(x_1), Q(x_2)]$	Nonlinear attenuation
	$min[Q(x_1), Q(x_2)] < Q(x_1 \cap x_2) < max[Q(x_1), Q(x_2)]$	Nonlinear attenuation（single factor）
	$Q(x_1 \cap x_2) > max[Q(x_1), Q(x_2)]$	Enhancement（two factors）
	$Q(x_1 \cap x_2) = Q(x_1) + Q(x_2)$	Independent
	$Q(x_1 \cap x_2) > Q(x_1) + Q(x_2)$	Nonlinear enhancement

among different tastes. In view of that, we adopted both the risk detector and interaction detector of the Geodector; the former was used for the significance evaluation of differences among death rates of CRC in our selected first-level administrative districts, while the latter was used for the detection of interaction between two tastes (x_1, x_2) (Table 1) and its influence on CRC.

V. DATA

A. DATA COLLECTION

1) COLORECTAL CANCER (CRC)

The death data used in this study come from the ‘Atlas of major diseases and causes of death in China’ published in Oct. 2016 by the National Centre for Chronic Noncommunicable Diseases Prevention and Control (http://ncncd.chinacdc.cn/),

TABLE 2. Cuisine and classification.

Data sources	Classification	Cuisine
Amap	Chinese restaurant	Sichuan, Hunan, Cantonese, Shanghai, Shandong, northeast, Jiangzhe, Beijing, Yunnan, Taiwan, Hubei, northwest, Anhui, Fujian, Chaozhou, Muslim cooking
Meishijie	Chinese cuisine	Sichuan, Hunan, Cantonese, northeast, Shandong, Zhejiang, Jiangsu, Halal, Fujian, Shanghai, Beijing, Hubei, Hui, Henan, northwest, Yungui, Jiangxi, Shanxi, Guangxi, Hong Kong, Taiwan, and other dishes
Meishichina	Chinese cuisine	Sichuan, Shandong, Fujian, Cantonese, Jiangsu, Zhejiang, Hunan, Anhui, Huaiyang, Henan, Jiangxi, Shanxi, Hubei, Yunnan, Beijing, Northeast, Northwest, Guizhou, Shanghai, Xinjiang, Hakka, Taiwan, Hong Kong, Macao
Haochu	Chinese eight Cuisines	Sichuan, Shandong, Hunan, Fujian, Jiangsu, Huizhou, Zhejiang, Cantonese
	Chinese local cuisine	Northeast, Chaozhou, Hubei, Yunnan, Guangxi, Jiangxi, Huaiyang, Northwest, Muslim cooking, Xinjiang, Beijing

which is affiliated with the Chinese Centre for Disease Control and Prevention (China CDC). Under the supervision of the National Health Commission, the China CDC is a governmental and national-level technical organization specializing in disease control and prevention as well as public health. It is also the official provider of disease-related data in China [46]. Moreover, all of the data provided by the China CDC have been stripped of information that may reveal patients' personal information. The atlas mentioned above records deaths for CRC per 0.1 million people in 33 Chinese first-level administrative districts (no data available in Taiwan Province), indicating the hazard that CRC has brought to public health [47].

2) POINTS OF INTEREST (POI) OF RESTAURANTS

For most restaurants, their menus must be designed after fully considering the dietary preferences of the locals, which implies that menus can reflect the local taste preference very well. As for the restaurant menus, the detailed information can be extracted from POI of restaurants.

AutoNavi is a digital map solution provider of content navigation and location-based services. Its primary product is Amap (https://www.amap.com/). Utilizing various means of data collection, such as high-precision aerial photography, vehicles, and walking, Amap stands out for its abundant data and clear documents. More importantly, Amap supports real-time data updates, which guarantees the comprehensiveness, accuracy and depth of its data [48].

In addition, compared with other digital maps, Amap classifies the POI of restaurants more meticulously. Those POI data offer detailed information about dishes served by each restaurant, including their names and the cuisines they are attributed to as well as the position coordinate of corresponding restaurant. Thus, the POI from AMaps have good indicator of reference. In this study, POI of restaurants were retrieved from Amap by our custom crawler in Aug. 2021. Then, the number of POI of each cuisine in each first-level administrative district was counted based on the fetched data. The classification of cuisine is presented in Table 2.

3) CROWDSOURCING CUISINE WEBSITE

Because Amap has more classifications of POI of restaurants than any single crowdsourcing cuisine website, to maintain consistency with its classification, we combined cuisine taste data from three different websites of a certain scale: Meishichina (https://www.meishichina.com/), Meishije (http://www.meishij.net/) and Haochu (https://www.haochu.com/).

As one of the largest Chinese gourmet websites, Meishichina provides users with over 0.1 million recipes detailing both major and minor ingredients as well as condiments. It is an important data source for food lovers as well as relevant researchers [55]. Meishijie, a food information service platform founded in Jan. 2007, successfully integrates elements of recipes, cooking skills, regional snacks, health tips, e-commerce and light social media. Its humanized design of content and searching has further promoted its popularity in China [56]. Haochu is a website that plays a role similar to a professional recipe dictionary, storing recipes for various dishes, nutrient information about ingredients, and health knowledge [57].

According to the birthplace of each dish, the three websites above sort Chinese dishes into several cuisines, such as Sichuan cuisine, Cantonese cuisine, and Northeastern Chinese cuisine (Table 2). It is worth noting that Henan cuisine

TABLE 3. The definition of each selected taste.

Taste	Definition	Reference	Representative Ingredient
Spicy	Hot feeling (mostly on the tongue) which is caused by capsicum et al.	[49]	Chilly
Pungent	Burning sensation in tongue, mouth, nose, which caused by something except for capsicum	[49]	Garlic, Ginger,
Sour	The taste sensation produced by acid	[50]	Vinegar
Sweet	The taste sensation typically produced by sugars (such as sucrose or glucose et al.)	[51]	Honey, Brown sugar
Umami	The taste sensation produced by several amino acids and nucleotides (such as glutamate and aspartate)	[52]	Chicken powder Aginomoto
Salty	The taste sensation produced by salt	[53]	Salt, Soy sauce
Oily	The taste sensation produced by fat	[54]	Cooking oil

(Yu cuisine) is only categorized as a separate cuisine in Meishijie and Meishichina, and there are no dishes under the Yu cuisine in Meishijie. The main reason for this is that Yu Cuisine is often regarded as the mother cuisine of all Chinese cuisines [58], [59]. Through comparisons among recipes of the same cuisine, we noticed clear differences in the frequency of utilization. Normally, a higher frequency leads to a higher impact on CRC. Therefore, we ranked all of the collected recipes by popularity and then retrieved their information, such as dish name and ingredients.

B. DATA PREPROCESSING

As Figure 2 shows, we collect POI data from three crowd-sourcing cuisine websites and Amap by crawlers. Between collection and analysis, there is also a data pre-processing aimed at getting the lists of Cuisine-Taste, Region-Cuisine and Region-Taste.

1) CUISINE_TASTE

Given that we collect recipes from three websites, data duplication should be addressed. Besides, since there is usually no strict limit on the dosage of ingredients such as spice or seasoning in Chinese cuisine, the lack of quantitative information about those ingredients is also a problem to be solved. If we want to figure out the relation between recipe and taste, we have to conduct quantization on collected data about ingredients. Therefore, with the integrated the recipe data, we set five hundred as the standardized amount of our collected recipes for each cuisine with reference to relevant researches and consideration of data characteristics. Next, we performed quantitative processing on ingredients based on their usage frequency in recipes gathered. The specific procedure is as follows.

The first step is data deduplication. We removed all the duplicate recipes including those only differing in names to count the recipe amount of each cuisine. Then we incorporated ingredients with synonyms. It is necessary to point out that we chose Meishichina as our main data resource. That is to say, for each cuisine, recipes from Meishijie and Haochu would be adopted in turn only when the number of the recipes published in meishichia is lower than five hundred.

The second step is to determine the research dimension of taste. Based on the comprehensive consideration of Chinese culinary culture [60], [61] eating habits [62] and previous studies on taste science [63], [64], we selected seven fundamental tastes (Spicy, Pungent, Sweet, Salty, Sour, Umami, and Oily). The selection of ‘Spicy’ and ‘Pungent’ came from the fact that, in China, people usually regard ‘Spicy’ and ‘Pungent’ as primary tastes instead of flavors. ‘Spicy’ specifically refers to taste stimulated only by capsicum, while ‘Pungent’ refers to taste stimulated by ingredients except for capsicum, such as garlic and ginger. The definition of each taste is given in Table 3. Utilizing the data related to ingredient taste, we calculated the matrix Ingredient_Taste.

The third step is labelling the taste of each dish. According to the matrix of Ingredient-Taste mentioned above, we labeled every spice and seasoning used in each dish. An ingredient may be labelled with several different tastes, for instance, cinnamon was labelled with ‘sweet’ and ‘pungent’ at the same time. And if different ingredients with the same taste label are used in one dish, that dish would be labelled with the taste several times.

After labelling, we counted the occurrence of each taste label to get the usage frequency of the seven tastes for every dish. Then based on their cuisine-belonging, we calculated the usage frequency for each cuisine to obtain the quantitative index of the seven tastes in Chinese cuisines. Based on the index, we produced the list of Cuisine-Taste.

2) REGION_CUISINE

This study collected POI data from Amap for around 603,467 restaurants nationwide. As mentioned earlier, the data can provide information about each restaurant like the address it locates and the cuisine it serves. Through matrix calculations on the POI data, this study constructed a Region_Cuisine matrix for 17 cuisines across 33 provincial-level administrative regions.

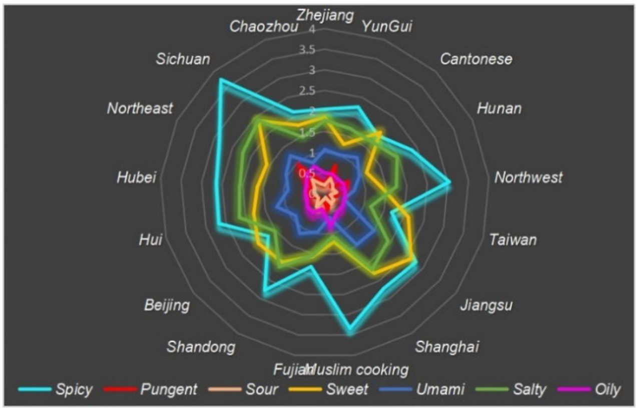

FIGURE 3. Visualization of the indexes of the seven tastes in the seventeen chinese cuisines.

3) REGION_TASTE

Based on the Region-Cuisine matrix, significant disparities are evident in the number of restaurant POI across study districts; for instance, Shandong's count (51,933) was over 196 times greater than Macao's (264). In addition, for each study district, the number of POI of different cuisine also varied a lot. Take Henan province as example, there were 9903 Muslim restaurants marked in Amap, while only 98 Shandong cuisine. Hence, before creating the list of Region-Taste, we conducted data rectification [24], [65]. The adopted rectification formula is defined as:

$$\text{Adjusting}_{\text{Taste}_{i,j_k}} = \frac{\text{Count}_{\text{POI}_{i,j}}}{\sum_{j=1}^{17} \text{Count}_{\text{POI}_{i,j}}} \text{Taste}_{i,j_k} \tag{6}$$

where Taste_{i,j_k} refers to kth taste of cuisine j in study district i, and $\text{Count}_{\text{POI}_{i,j}}$ represents the number of restaurants of cuisine j in study district i, while $\text{Adjusting\_Taste}_{i,j_k}$ denotes the rectified kth taste of cuisine j in study district i.

VI. RESULTS

A. TASTE CHARACTERISTICS OF CUISINES

Through labelling and calculation, we obtained the characteristics of the seven fundamental tastes in the seventeen Chinese cuisines. The results are presented in Table 4 and Figure 3. Generally, among the seven tastes, 'Spicy' was the most often used in our study cuisines, followed by 'Sweet' and 'Salty'. However, 'Pungent' and 'Sour' were the least frequently used tastes (Figure 3). Specifically, Sichuan cuisine was characterized by a heavy taste, judging from its high usage frequency of all seven tastes. In contrast, Taiwan cuisine was much lighter, with its focus on the original flavour of ingredients. In addition, Beijing cuisine was often very sweet but rarely spicy, whereas Northwest cuisine and Muslim food were in the opposite style. Cantonese cuisine and Zhejiang cuisine usually emphasized the taste of 'Sweet' and 'Umami'. The taste usage in Hubei cuisine was relatively balanced.

	Hui	Beijing	Chaozhou	Northeast	Fujian	Cantonese	Hubei	Hunan	Sauerkraut	Muslim cooking	Shandong	Shanghai	Sichuan	Northwest	Yungui	Zhejiang	Taiwan
Macao	0.00	0.00	5.88	7.35	0.00	47.06	0.00	1.47	0.00	1.47	0.00	17.65	7.35	0.00	1.47	0.00	10.29
Hong Kong	0.17	1.67	21.70	0.00	0.17	39.57	0.00	1.84	0.33	0.83	0.00	10.68	9.85	0.00	4.51	0.67	8.01
Qinghai	0.05	0.11	0.00	1.02	0.03	0.75	0.05	2.48	0.11	70.53	0.19	0.13	20.28	3.45	0.45	0.35	0.03
Chongqing	0.13	3.37	1.28	3.30	0.47	6.67	0.27	4.72	0.67	11.59	0.27	0.20	60.31	1.28	3.23	1.68	0.54
Tibet	0.25	0.33	0.08	1.64	0.16	0.82	0.16	2.71	0.16	20.81	0.49	0.41	57.15	13.08	1.32	0.41	0.00
Guizhou	0.10	0.86	0.15	1.06	0.33	2.15	0.08	7.44	0.23	32.07	0.46	0.23	30.48	0.35	23.32	0.53	0.18
Ningxia	0.08	0.23	0.02	0.93	0.00	0.40	0.02	0.42	0.15	86.39	0.17	0.04	8.06	2.61	0.28	0.17	0.02
Sichuan	0.05	1.05	0.29	1.41	0.23	2.65	0.13	1.16	0.41	18.11	0.35	0.24	69.20	1.85	1.60	0.96	0.32
Gansu	0.08	0.16	0.01	0.82	0.02	0.52	0.02	1.59	0.11	69.46	0.15	0.05	23.41	2.95	0.44	0.18	0.02
Hainan	0.67	0.76	2.00	6.10	0.48	11.71	0.38	22.48	0.19	19.24	0.38	0.19	28.57	2.67	2.95	0.57	0.67
Guangxi	0.31	0.96	2.91	4.59	0.80	27.07	0.34	20.84	0.69	11.20	0.76	0.31	23.06	1.41	1.64	1.53	1.57
Hunan	0.15	0.63	0.34	1.29	0.14	2.43	0.15	70.25	0.28	14.38	0.26	0.48	7.36	0.45	0.68	0.55	0.18
Xinjiang	0.17	0.23	0.03	1.06	0.07	0.52	0.07	2.97	0.14	66.79	0.40	0.14	24.47	2.27	0.27	0.34	0.04
Hubei	0.17	0.78	0.46	1.55	0.16	1.98	26.18	15.09	0.44	23.67	0.16	0.13	27.08	0.51	0.53	0.78	0.34
Yunnan	0.15	0.31	0.16	0.94	0.13	1.29	0.14	2.07	0.17	50.73	0.24	0.17	21.01	0.44	21.53	0.34	0.21
Inner Mongolia	0.44	0.79	0.09	14.07	0.11	1.18	0.09	1.84	0.24	58.03	1.03	0.15	12.78	5.70	2.32	1.09	0.07
Shanxi	0.68	1.44	0.07	6.67	0.14	2.38	0.21	7.40	0.42	36.63	0.73	0.31	35.69	3.82	1.89	1.18	0.33
Jiangxi	1.05	1.67	1.40	2.37	0.94	5.76	0.35	29.44	0.77	28.99	0.59	0.59	22.39	0.63	1.33	1.40	0.35
Fujian	0.50	1.13	1.17	2.43	21.51	4.91	0.36	8.00	0.25	17.04	0.40	0.22	37.46	0.70	0.82	1.47	1.62
Shaanxi	0.46	0.93	0.20	3.52	0.24	1.91	0.22	9.22	0.37	31.13	0.47	0.22	39.43	9.88	1.05	0.63	0.14
Henan	0.33	0.95	0.12	2.43	0.20	1.87	1.23	5.28	0.18	64.58	0.72	0.23	18.59	1.79	0.73	0.59	0.18
Guangdong	0.24	0.46	11.74	2.18	0.28	26.61	1.14	18.69	0.20	13.89	0.25	0.13	20.44	1.59	0.88	0.51	0.79
Jilin	2.49	0.74	0.16	24.86	0.11	1.97	0.14	1.59	0.14	40.69	0.60	0.11	22.48	0.47	2.74	0.52	0.19
Tianjin	2.72	1.53	0.19	8.58	0.30	2.53	0.22	3.36	0.40	57.41	1.99	0.43	15.55	1.83	1.29	0.83	0.83
Liaoning	2.06	1.49	0.28	22.39	0.11	1.93	0.16	1.42	0.31	41.56	0.80	0.33	22.19	1.12	2.48	0.88	0.49
Beijing	3.61	14.03	0.91	11.71	0.36	6.24	0.91	10.95	0.73	18.09	1.62	0.61	18.09	4.93	4.44	1.60	1.17
Heilongjiang	6.78	1.02	0.29	26.10	0.17	1.39	0.17	1.20	0.27	42.02	0.93	0.20	15.80	0.80	1.76	0.54	0.56
Shandong	2.00	1.16	0.09	6.89	0.15	1.25	0.10	1.50	0.32	40.68	25.10	0.16	17.61	1.18	1.19	0.48	0.15
Shanghai	9.20	1.50	1.48	5.18	0.68	9.36	0.16	13.23	1.30	15.97	1.10	14.38	15.90	2.92	1.69	3.61	2.35
Zhejiang	3.37	1.19	0.37	2.55	0.37	2.29	0.15	6.31	0.24	25.19	0.36	0.27	34.80	1.12	1.37	19.56	0.46
Jiangsu	4.34	1.49	0.39	3.48	0.30	2.56	0.20	5.97	14.04	28.56	0.98	0.99	28.01	2.05	1.41	4.48	0.77
Hebei	5.99	2.05	0.02	9.55	0.10	0.95	0.16	2.91	0.20	50.13	1.00	0.16	22.26	1.83	1.84	0.69	0.14
Anhui	47.89	0.87	0.16	1.71	0.27	1.07	0.09	3.31	0.45	24.26	0.39	0.41	15.87	0.65	0.93	1.44	0.23

0 86

FIGURE 4. Proportions of various types of POI in the 33 study districts.

B. SPATIAL DISTRIBUTION OF TASTES

The restaurant POI data fetched in this study covered 33 provincial administrative districts and seventeen study cuisines. Through data analysis, we noticed that, for most cuisines, majority of their corresponding restaurants were located in their birthplaces. Compared with other cuisines, Muslim restaurants and Sichuan restaurants occupied larger market shares outside their places of origin, while Beijing restaurants and Taiwan restaurants took the smallest market shares in most districts (Figure 4). Among the 33 first-level administrative districts, Chinese restaurants in Macao covered the fewest kinds of cuisines (9), with Hong Kong coming in a close second (13). Specially, we also noted that most Muslim restaurants were located in Henan rather than Ningxia, Xinjiang, or Gansu, where more Muslims populated. A similar situation also gripped Hunan restaurants and Sichuan restaurants, which mostly opened in Guangdong, not in their birthplaces. This shows the inclusiveness and diversity of Guangdong food culture [66].

Considering the differences among the absolute values of taste usage in different districts (sometimes even in the same district), we standardized the data referring to the existing relevant studies and mapped the spatial distribution of the tastes, as shown in Figure 5. Overall, food served in northern China was often relatively 'Oily', while the southern region preferred tastes such as 'Spicy', 'Sweet', 'Umami' and

TABLE 4. Indexes of the seven tastes in the seventeen chinese cuisines.

Cuisine	Taste						
	Spicy	Pungent	Sour	Sweet	Umami	Salty	Oily
Zhejiang	2.0436	0.2436	0.2040	1.8455	1.0634	1.7723	0.4970
Yungui	2.2477	0.7207	0.3874	1.2793	1.0000	1.6892	0.5000
Cantonese	1.9024	0.2151	0.2849	2.0000	1.1713	1.6952	0.4701
Hunan	2.3591	0.6528	0.2242	1.1329	0.9385	1.9623	0.5139
Northwest	3.0360	0.5540	0.2881	1.4017	0.7147	1.7452	0.4958
Taiwan	1.6883	0.2764	0.1436	2.1084	0.4851	1.1626	0.5014
Jiangsu	2.7717	0.1196	0.2717	2.6304	1.5000	1.9239	0.5326
Shanghai	2.7164	0.1940	0.2985	2.2836	1.4627	2.1791	0.5522
Muslim cooking	3.3333	0.4000	0.1000	1.2000	0.6000	1.0667	0.8000
Fujian	1.8078	0.1631	0.1631	1.5612	0.9592	1.5340	0.3883
Shandong	2.7649	0.3386	0.3725	1.9641	1.1474	2.0976	0.5478
Beijing	1.7206	0.1580	0.1927	2.0173	0.8902	1.5164	0.5318
Hui	2.6714	0.3714	0.2057	1.7914	1.2086	2.1514	0.4429
Hubei	2.6432	0.3970	0.2563	1.6432	0.8945	2.0653	0.4774
Northeast	2.8558	0.3450	0.2710	1.5731	1.0448	2.2203	0.4912
Sichuan	3.7401	0.9821	0.4464	2.3810	1.2341	2.3909	0.6389
Chaozhou	2.1197	0.3451	0.2465	1.7746	0.7958	1.4648	0.7042

'Salty'. From a local view, 'Spicy' prevailed in the northwester and southwestern regions. Moreover, Figure 6 shows the high usage frequencies of seven tastes, except 'Oily', even including 'Umami' in Sichuan and Chongqing. The observed frequent usage of 'Umami' in these two southwestern cities was somewhat different from the common perception [67]. In addition, Figure 6 suggests that foods served in Shanghai, Zhejiang and Fujian tended to be on the sweet side. Those served in Guangxi, Guangdong, Fujian, Hainan, Hong Kong, Macao, Anhui had the lowest levels of oil [62].

C. RELATIVE IMPORTANCE OF TASTES

Figure 7 shows the LMG model result. As seen in this figure, the three tastes with the highest relative importance were 'Spicy' (31.38%), 'Pungent' (21.76%), and 'Oily' (13.15%), which implies that they may generate a greater influence on CRC than other tastes. However, 'Pungent' and 'Oily' were used least often in all seventeen cuisines (Figure 3). Moreover, Figure 7 shows that the tastes with the lowest relative importance were 'Sour' (7.92%) and 'Salty' (5.31%). Figure 3 shows that 'Salty' had the second-highest usage frequency.

D. RISK DETECTION OF TASTES

The risk detection result of tastes is shown in Figure 8. This figure is plotted by the standardized contribution value of tastes on the horizontal axis (the contribution value of a taste is proportional to its usage frequency) and the average mortality rate of CRC in the corresponding study district in 2015 on the vertical. Regarding the most important tastes, namely 'Pungent' and 'Spicy', there were clear inverse relationships between their contribution values within a certain range and the corresponding mortality rate (Figure 8). The lowest contribution values of both of them corresponded to the highest mortality. When their contribution values were -0.022 (Pungent) and 1.205 (Spicy), the lowest death rates were observed and their overall trends were falling despite local fluctuation.

TABLE 5. The results of the interaction relationship.

	Spicy	Pungent	Sour	Sweet	Umami	Salty	Oily
Spicy	0.8422**						
Pungent	0.9188	0.8080**					
Sour	0.8804	0.9122	0.1670				
Sweet	0.9151	0.9228	0.3358	0.2919			
Umami	0.9672	0.9823	0.5594	0.6247	0.4885*		
Salty	0.9050	0.9224	0.5882	0.8883	0.9229	0.1744	
Oily	0.9615	0.9276	0.6066	0.9612	0.8033	0.7001	0.6559**

Note: "*" and "**" indicate p ≤ 0.05 and p ≤ 0.01 in the F-test, respectively.

E. INTERACTION OF TASTES

At the level of $p < 0.01$, 'Spicy', 'Pungent' and 'Oily' were significantly positively associated with CRC mortality (Table 5), which is consistent with the LMG model result. Hence, these three tastes can be regarded as the top three important taste risk factors for CRC.

All of the interactions among the seven tastes are listed in Table 6. According to Table 6, each interaction involved

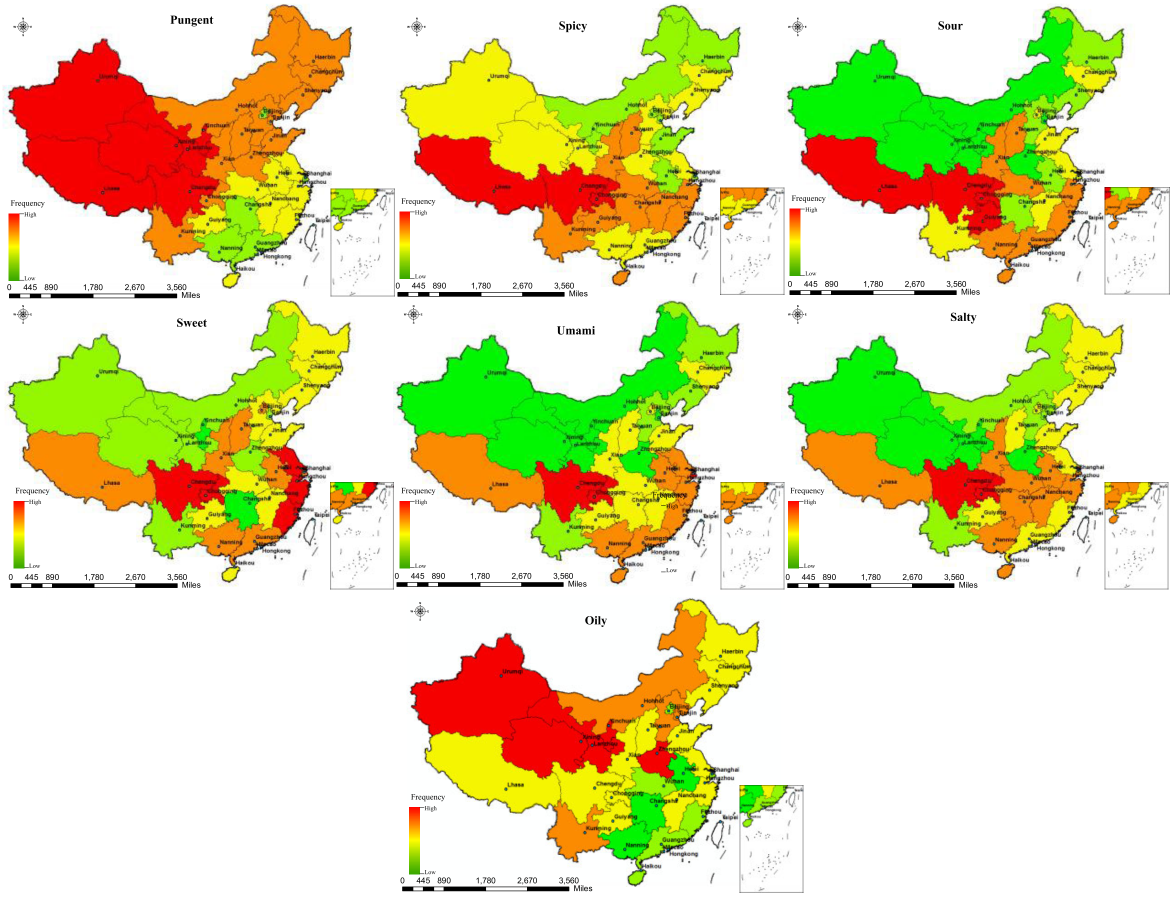

FIGURE 5. Spatial distribution of the seven tastes.

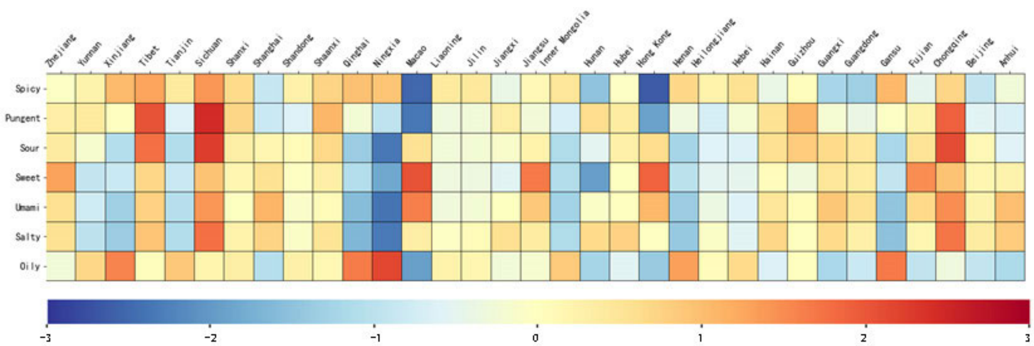

FIGURE 6. Standardized indexes of the seven tastes in provincial regions.

with 'Spicy' or 'Pungent' correlated with increased CRC mortality to some degree. Also, the interactions of 'Salty' & 'Sour', 'Salty' & 'Sweet', 'Salty' & 'Umami', 'Salty' & 'Oily', 'Oily' & 'Sour', 'Oily' & 'Sweet' and 'Oily' & 'Salty' all led to significant nonlinear enhancement effects on CRC risk. However, 'Salty' and 'Sour' were ranked poorly in the relative importance assessment of a single taste (Figure 7).

VII. DISCUSSION

The close link between CRC and dietary habits is widely acknowledged in the medical community, and taste preference is an important component of dietary habits [68], [69]. There is so much nutrition information hidden behind people's taste preferences because a particular taste preference usually indicates cravings for the nutrients associated with the taste [70], [71]. Differences in factors, such as geographical location and climate, make different tastes present across the country. Thus, differences in taste presence in some way result in slight differences in various nutrients received by humans in different regions [72], [73]. In addition, individual perception of taste relies on an individual taste system, which can recognize and quantify nutrients behind

TABLE 6. The results of the interaction relationship.

	Spicy	Pungent	Sour	Sweet	Umami	Salty
Pungent	Enhanced, bi-					
Sour	Enhanced, bi-	Enhanced, bi-				
Sweet	Enhanced, bi-	Enhanced, bi-	Enhanced, bi-			
Umami	Enhanced, bi-	Enhanced, bi-	Enhanced, bi-	Enhanced, bi-		
Salty	Enhanced, bi-	Enhanced, bi-	Enhance, nonlinear	Enhance, nonlinear	Enhance, nonlinear	
Oily	Enhanced, bi-	Enhanced, bi-	Enhance, nonlinear	Enhance, nonlinear	Enhanced, bi-	Enhance, nonlinear

Note: Enhance, nonlinear; Enhanced, bi-

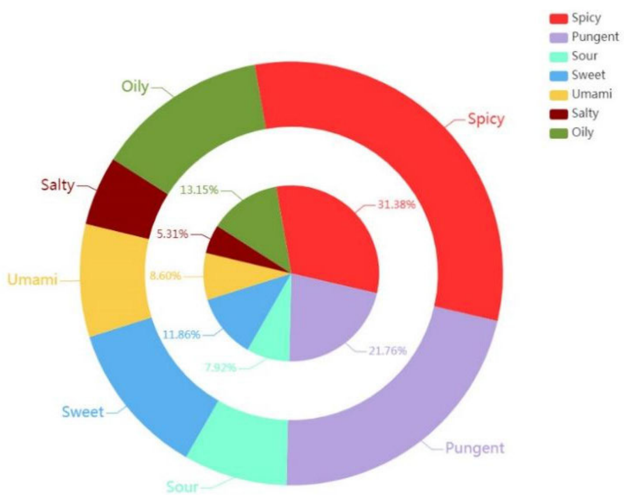

FIGURE 7. The relative importance of tastes in their influences on CRC.

the taste [10], [47]. When someone's taste system is disturbed by something like disease, his or her appetite for food as well as taste choice may be affected.

Recent studies have indicated that the exploration of taste breaks fresh ground for the prevention and detection of chronic disease [47], [74]. And several researchers have confirmed the association between taste and chronic disease in biology [75], [76], genetics [77], [78], [79], and nutriology [80], [81]. Their researches provide theoretical support and technical references for further study. The study presented in this paper is grounded in the potential quantitative relationship between taste preference and CRC. From the perspective of the spatial patterns of CRC and taste, the spatial distribution characteristics of the two within first-level Chinese administrative districts was explored to find their spatial relationship. Furthermore, the relative importance of various food tastes to CRC was analyzed and the quantitative relationships among taste risk factors for CRC were mined. The study method proposed for this study is suitable and referenced for other relevant studies, and the results we have obtained present statistical significance.

Despite the significance mentioned above, as an exploratory study, there is something worthy of discussion:

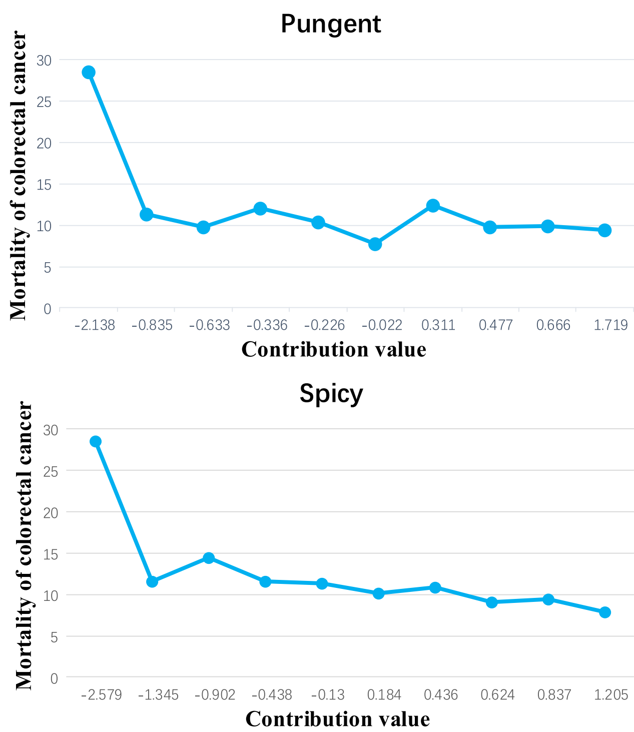

FIGURE 8. Results of risk detector of pungent and spicy tastes for CRC.

(1) Data. Three issues we need to deal when it came to data. First is incomplete. The restaurant POI data we used was from 2021 Amap. But it was incomplete because of various limitaions. Besides, there was significant data overlap, since the recipe data came from three different crowdsourcing cuisine websites. Although we have performed data pretreatment, noisy data still existed due to data size and cognitive differences. Last but not least, on account of data availability issues, the latest CRC death data we could use were statistics for 2016 from the Chinese Centre for Disease Control and Prevention, whereas the data about restaurants and recipes were from 2021. We must admit failure to achieve an optimal logic with respect to the temporal design, and it is undeniable that this is one of the limitations of this study. Even so, considering that eating habits are formed over a long period of time and normally will not change significantly within a

short period, the scientific validity of the taste characteristics obtained could be assured [1], [4], [5], [9], [14], [16], [71], [79], [80], [82].

(2) Methodology. From the perspective of the spatial patterns of CRC and taste and within large-scale, first-level administrative districts, we evaluated the relative importance of different tastes to CRC by the LMG model and utilized Geodector to establish a spatial relationship. These methods contribute to the strengths and weaknesses of our study. On the one hand, the ability to detect various relationships, including interactive relationships and linear and nonlinear relationships, puts Geodector at an advantage when compared with other bivariate regression models. On the other hand, both the LMG model and Geodector rely on spatial statistics, however, our statistics have not been verified by rigorous clinical trials, which affects the credibility of the study results to some degree.

(3) Results. Considering some findings are proposed for the first time, their rationalities remain to be confirmed by clinical trials in disciplines such as biology pathology. However, it has to be acknowledged that, due to the data-driven analysis, this paper may not have sufficiently addressed ecological and other factors, potentially leading to the following limitations in the findings: ① Incomplete understanding of environmental Risk Factors. Important ecological contributors to CRC may have been overlooked; ② Limited universality. The research results may not be generalizability and applied universally across different countries; ③ Single perspective. This study may not fully explain the multifactorial nature of CRC.

Furthermore, due to the reasons mentioned earlier, Yu cuisine has not been incorporated into this study. However, through its long history of evolution, Yu cuisine has cultivated unique culinary techniques, ingredient combinations, and dietary culture. Not considering these elements might lead to the omission of significant taste preference risk factors, potentially compromising the comprehensiveness of the study results. Future studies from our team may need to take these factors into account to achieve more persuasive conclusions.

VIII. CONCLUSION

This paper presents an exploratory quantitative study on the spatial relationship between CRC and taste using crowdsourcing data. Statistically significant results and several new findings are obtained: (1) Chinese taste preference presented significant spatial heterogeneity, and the most often used tastes in Chinese cuisines were 'Spicy', 'Sweet', and 'Salty'; (2) 'Spicy' and 'Pungent' were the top taste risk factors for CRC, with the highest relative importance; (3) the taste interactions involved with 'Salty' or 'Oily' might increase CRC risk.

Over and above the findings, there are two limitations of this study. On the one hand, the deficiency of optimal logic design and the inadequacy of attention to clinical trials have more or less affected the rigor of this study. On the other hand, regarding the study results, the study's explanation of environmental and ecological factors is not comprehensive enough, and the research perspective is somewhat limited, which may lead to limited generalizability in the conclusions.

REFERENCES

[1] M. Arnold, M. S. Sierra, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global patterns and trends in colorectal cancer incidence and mortality," *Gut*, vol. 66, no. 4, pp. 683–691, Apr. 2017.

[2] S. Fukuoka, H. Hara, N. Takahashi, T. Kojima, A. Kawazoe, M. Asayama, T. Yoshii, D. Kotani, H. Tamura, Y. Mikamoto, N. Hirano, M. Wakabayashi, S. Nomura, A. Sato, T. Kuwata, Y. Togashi, H. Nishikawa, and K. Shitara, "Regorafenib plus nivolumab in patients with advanced gastric or colorectal cancer: An open-label, dose-escalation, and dose-expansion phase ib trial (REGONIVO, EPOC1603)," *J. Clin. Oncol.*, vol. 38, no. 18, pp. 2053–2061, Jun. 2020.

[3] R. L. Siegel, K. D. Miller, A. G. Sauer, S. A. Fedewa, L. F. Butterly, J. C. Anderson, A. Cercek, R. A. Smith, and A. Jemal, "Colorectal cancer statistics, 2020," *CA, A Cancer J. Clinicians*, vol. 70, no. 3, pp. 145–164, May 2020.

[4] Y.-H. Xie, Y.-X. Chen, and J.-Y. Fang, "Comprehensive review of targeted therapy for colorectal cancer," *Signal Transduction Targeted Therapy*, vol. 5, no. 1, p. 30, Mar. 2020.

[5] M. Blankenburg, M. Elhamamy, D. Zhang, N. Fujikawa, A. Corbin, G. Jin, J. Harris, and G. Knobloch, "Evaluation of health economic impact of initial diagnostic modality selection for colorectal cancer liver metastases in suspected patients in China, Japan and the USA," *J. Med. Econ.*, vol. 26, no. 1, pp. 219–232, Dec. 2023.

[6] R. L. Siegel, K. D. Miller, S. A. Fedewa, D. J. Ahnen, R. G. S. Meester, A. Barzi, and A. Jemal, "Colorectal cancer statistics, 2017," *CA, A Cancer J. Clinicians*, vol. 67, no. 3, pp. 177–193, May 2017.

[7] R. Feng, Y. Zong, S. Cao, and R. Xu, "Current cancer situation in China: Good or bad news from the 2018 global cancer statistics?" *Cancer Commun.*, vol. 39, no. 1, pp. 1–12, Dec. 2019.

[8] M. Song, A. T. Chan, and J. Sun, "Influence of the gut microbiome, diet, and environment on risk of colorectal cancer," *Gastroenterology*, vol. 158, no. 2, pp. 322–340, Jan. 2020.

[9] N. Bhimani, G. Y. Wong, C. Molloy, M. Dieng, and T. J. Hugh, "Cost of colorectal cancer by treatment type from different health economic perspectives: A systematic review," *Eur. J. Surgical Oncol.*, vol. 48, no. 10, pp. 2082–2093, Oct. 2022.

[10] K. E. Bradbury, N. Murphy, and T. J. Key, "Diet and colorectal cancer in U.K. Biobank: A prospective study," *Int. J. Epidemiol.*, vol. 49, no. 1, pp. 246–258, Feb. 2020.

[11] Y. Meng, S. Li, D. Gu, K. Xu, M. Du, L. Zhu, H. Chu, Z. Zhang, Y. Wu, Z. Fu, and M. Wang, "Genetic variants in m^6A modification genes are associated with colorectal cancer risk," *Carcinogenesis*, vol. 41, no. 1, pp. 8–17, Oct. 2019.

[12] P. Favoriti, G. Carbone, M. Greco, F. Pirozzi, R. E. M. Pirozzi, and F. Corcione, "Worldwide burden of colorectal cancer: A review," *Updates Surgery*, vol. 68, no. 1, pp. 7–11, Mar. 2016.

[13] X. He, K. Wu, X. Zhang, R. Nishihara, Y. Cao, C. S. Fuchs, E. L. Giovannucci, S. Ogino, A. T. Chan, and M. Song, "Dietary intake of fiber, whole grains and risk of colorectal cancer: An updated analysis according to food sources, tumor location and molecular subtypes in two large U.S. Cohorts," *Int. J. Cancer*, vol. 145, no. 11, pp. 3040–3051, Dec. 2019.

[14] M. J. Sneyd and B. Cox, "Do low-fat foods alter risk of colorectal cancer from processed meat?" *Public Health*, vol. 183, pp. 138–145, Jun. 2020.

[15] L. Schwingshackl, C. Schwedhelm, G. Hoffmann, S. Knüppel, A. L. Preterre, K. Iqbal, A. Bechthold, S. De Henauw, N. Michels, B. Devleesschauwer, H. Boeing, and S. Schlesinger, "Food groups and risk of colorectal cancer," *Int. J. Cancer*, vol. 142, no. 9, pp. 1748–1758, May 2018.

[16] M. Song, W. S. Garrett, and A. T. Chan, "Nutrients, foods, and colorectal cancer prevention," *Gastroenterology*, vol. 148, no. 6, pp. 1244–1260, May 2015.

[17] E. M. Mahfouz, R. R. Sadek, W. M. Abdel-Latief, F. A.-H. Mosallem, and E. E. Hassan, "The role of dietary and lifestyle factors in the development of colorectal cancer: Case control study in Minia, Egypt," *Central Eur. J. Public Health*, vol. 22, no. 4, pp. 215–222, Dec. 2014.

[18] K. Almurshed and R. Nashar, "Colorectal cancer: A case control study of dietary factors, King Faisal specialist hospital and researh center, Riyadh, Saudi Arabia," *J. Family Community Med.*, vol. 15, no. 2, p. 57, 2008.
[19] S.-A. Lee, X. O. Shu, G. Yang, H. Li, Y.-T. Gao, and W. Zheng, "Animal origin foods and colorectal cancer risk: A report from the Shanghai Women's health study," *Nutrition Cancer*, vol. 61, no. 2, pp. 194–205, Mar. 2009.
[20] A. Ramadas and M. Kandiah, "Nutritional status and the risk for colorectal adenomas: A case-control study in hospital Kuala Lumpur, Malaysia," *Pakistan J. Nutrition*, vol. 9, no. 3, pp. 269–278, Feb. 2010.
[21] S. P. Nayak, M. Sasi, M. Sreejayan, and S. Mandal, "A case-control study of roles of diet in colorectal carcinoma in a South Indian Population," *Asian Pacific J. Cancer Prevention*, vol. 10, no. 4, pp. 565–568, 2009.
[22] S. Azeem, S. W. Gillani, A. Siddiqui, S. B. Jandrajupalli, V. Poh, and S. A. S. Sulaiman, "Diet and colorectal cancer risk in Asia—A systematic review," *Asian Pacific J. Cancer Prevention*, vol. 16, no. 13, pp. 5389–5396, Aug. 2015.
[23] X. Tian, G. Yu, and F. He, "An analysis of sleep complaints on Sina Weibo," *Comput. Hum. Behav.*, vol. 62, pp. 230–235, Sep. 2016.
[24] A. Hausmann, T. Toivonen, R. Slotow, H. Tenkanen, A. Moilanen, V. Heikinheimo, and E. D. Minin, "Social media data can be used to understand Tourists' preferences for nature-based experiences in protected areas," *Conservation Lett.*, vol. 11, no. 1, Jan. 2018, Art. no. e12343.
[25] C. L. Vecchia, S. Franceschi, E. Bidoli, F. Barbone, and P. Dolara, "Refined-sugar intake and the risk of colorectal cancer in humans," *Int. J. Cancer*, vol. 55, no. 3, pp. 386–389, Sep. 1993.
[26] M. Shen, L. Shi, L. Li, S. Zhang, C. Zhang, and Y. Jiang, "A case-control study on urinary calculi and dietary factors," *Chin. J. Epidemiol.*, vol. 23, no. 2, pp. 134–137, 2002.
[27] Y.-P. Su, K. Yang, and X. T. Liu, "Relationship between the dietary behavior and the risk of diabetes in adults," *Chin. J. Prevention Control Chronic Diseases*, vol. 7, pp. 481–485, Dec. 2018.
[28] H. Qin, Q. Huang, Z. Zhang, Y. Lu, M. Li, L. Xu, and Z. Chen, "Carbon dioxide emission driving factors analysis and policy implications of Chinese cities: Combining geographically weighted regression with two-step cluster," *Sci. Total Environ.*, vol. 684, pp. 413–424, Sep. 2019.
[29] R. H. Lindeman, P. F. Merenda, and R. Z. Gold, "Introduction to bivariate and multivariate analysis," *J. Amer. Stat. Assoc.*, vol. 76, no. 375, p. 752, 1980, doi: 10.2307/2287559.
[30] W. Kruskal, "Relative importance by averaging over orderings," *Amer. Statistician*, vol. 41, no. 1, p. 6, Feb. 1987.
[31] W. Kruskal, "Correction: Relative importance by averaging over orderings," *Amer. Statistician*, vol. 41, no. 4, p. 341, 1987.
[32] M. I. Idrissi, V. Chabridon, and B. Iooss, "Developments and applications of Shapley effects to reliability-oriented sensitivity analysis with correlated inputs," *Environ. Model. Softw.*, vol. 143, Sep. 2021, Art. no. 105115.
[33] J. Wang and C. Xu, "Geodetector: Principle and prospective," *Acta Geographica Sinica*, vol. 72, no. 1, pp. 116–134, 2017.
[34] J. Gu, L. Liang, H. Song, Y. Kong, R. Ma, Y. Hou, J. Zhao, J. Liu, N. He, and Y. Zhang, "A method for hand-foot-mouth disease prediction using GeoDetector and LSTM model in Guangxi, China," *Sci. Rep.*, vol. 9, no. 1, pp. 1–10, Nov. 2019.
[35] J. Li, J. Liang, J. Wang, Z. Ren, D. Yang, Y. Wang, Y. Mu, X. Li, M. Li, Y. Guo, and J. Zhu, "Spatiotemporal trends and ecological determinants in maternal mortality ratios in 2,205 Chinese counties, 2010–2013: A Bayesian modelling analysis," *PLoS Med.*, vol. 17, no. 5, 2020, Art. no. e1003114.
[36] L. Wang, C. Xu, J. Wang, J. Qiao, M. Yan, and Q. Zhu, "Spatiotemporal heterogeneity and its determinants of COVID-19 transmission in typical labor export provinces of China," *BMC Infectious Diseases*, vol. 21, no. 1, pp. 1–12, Dec. 2021.
[37] F. Pei, R. Zhong, L.-A. Liu, and Y. Qiao, "Decoupling the relationships between carbon footprint and economic growth within an urban agglomeration—A case study of the Yangtze River Delta in China," *Land*, vol. 10, no. 9, p. 923, Sep. 2021.
[38] L. Wang, C. Xu, M. Hu, J. Qiao, W. Chen, T. Li, S. Qian, and M. Yan, "Spatio-temporal variation in tuberculosis incidence and risk factors for the disease in a region of unbalanced socio-economic development," *BMC Public Health*, vol. 21, no. 1, pp. 1–11, Dec. 2021.
[39] Q. Xia, X. Zhang, Y. Hu, W. Tian, W. Miao, B. Wu, Y. Lai, J. Meng, Z. Fan, C. Zhang, L. Xin, J. Miao, Q. Wu, M. Jiao, L. Shan, N. Wang, B. Shi, and Y. Li, "The superposition effects of air pollution on government health expenditure in China—Spatial evidence from GeoDetector," *BMC Public Health*, vol. 22, no. 1, pp. 1–15, Dec. 2022.
[40] G. Rong, K. Li, L. Han, S. Alu, J. Zhang, and Y. Zhang, "Hazard mapping of the rainfall–landslides disaster chain based on GeoDetector and Bayesian network models in Shuicheng County, China," *Water*, vol. 12, no. 9, p. 2572, Sep. 2020.
[41] Y. Wang, H. Wen, D. Sun, and Y. Li, "Quantitative assessment of landslide risk based on susceptibility mapping using random forest and GeoDetector," *Remote Sens.*, vol. 13, no. 13, p. 2625, Jul. 2021.
[42] J. Lin, W. Chen, X. Qi, and H. Hou, "Risk assessment and its influencing factors analysis of geological hazards in typical mountain environment," *J. Cleaner Prod.*, vol. 309, Aug. 2021, Art. no. 127077.
[43] Y. Zhou, X. Li, and Y. Liu, "Land use change and driving factors in rural China during the period 1995–2015," *Land Use Policy*, vol. 99, Dec. 2020, Art. no. 105048.
[44] M. Chen, Z. Bai, Q. Wang, and Z. Shi, "Habitat quality effect and driving mechanism of land use transitions: A case study of Henan water source area of the middle route of the South-to-North water transfer project," *Land*, vol. 10, no. 8, p. 796, Jul. 2021.
[45] Y. Liu, K. Wu, and H. Cao, "Land-use change and its driving factors in Henan province from 1995 to 2015," *Arabian J. Geosci.*, vol. 15, no. 3, pp. 1–16, Feb. 2022.
[46] (Mar. 23, 2022). *Chinese Center for Disease Control and Prevention*. [Online]. Available: https://www.chinacdc.cn/en/
[47] H. Li, P. Jia, and T. Fei, "Associations between taste preferences and chronic diseases: A population-based exploratory study in China," *Public Health Nutrition*, vol. 24, no. 8, pp. 1–12, Jun. 2020.
[48] Z. Zheng, "Spatial distribution characteristics and influencing factors of occupation and residence in Zhuhai: Investigation and analysis based on interest points of Gaode map and traffic sensing data/Zheng Zeshuang," *Planners*, vol. 35, no. 4, pp. 51–56, 2019.
[49] H. Li, P. Jia, and T. Fei, "Geographical association between dietary tastes and chronic diseases in China: An exploratory study using crowdsourcing data mining techniques," *Acta Geographica Sinica*, vol. 74, no. 8, p. 1637, 2019.
[50] *Merriam-Webster*. [Online]. Available: https://www.merriam-webster.com/dictionary/sour
[51] Merriam-Webster. *Sweet*. [Online]. Available: https://www.merriam-webster.com/dictionary/sweet
[52] Merriam-Webster. *Umami*. [Online]. Available: https://www.merriam-webster.com/dictionary/umami
[53] Merriam-Webster. *Salty*. [Online]. Available: https://www.merriam-webster.com/dictionary/salty
[54] C. A. Running, B. A. Craig, and R. D. Mattes, "Oleogustus: The unique taste of fat," *Chem. Senses*, vol. 40, no. 7, pp. 507–516, Sep. 2015.
[55] W. Xu, Y. Li, L. Wei, and Z. Wu, "A feature-based food recommendation on apache spark," in *Proc. 5th Int. Conf. Adv. Cloud Big Data (CBD)*, Shanghai, China, Aug. 2017, pp. 123–128.
[56] B. Wikipedia. (2021). *Meishije*. [Online]. Available: https://baike.baidu.com/item/%E7%BE%8E%E9%A3%9F%E6%9D%B0/9631761?fr=aladdin
[57] Haochu. (2021). *Haochu*. [Online]. Available: https://www.haochu.com/
[58] X. Qiao, "A complex network study of the eight Chinese cuisines," M.S. thesis, School Phys. Astron., Yunnan Univ., Kunming, China, 2018.
[59] Z. Wang, "An analysis of metaphor-loaded dish names of Chinese eight major culinary schools from the perspective of conceptual metaphor," M.S. thesis, College Int. Stud., Southwest Univ., Chongqing, China, 2016.
[60] C. Chen, "The culture of Chinese diet: Regional differentiation and developing trends," *Acta Geographica Sinica*, vol. 49, no. 3, pp. 226–235, 1994.
[61] J. M. Newman, *Food Culture in China*. Westport, CT, USA: Greenwood Publishing Group, 2004.
[62] L. Yong, "The reasons and distribution of pungent flavour districts in China's dietetics," *Geographical Res.*, vol. 20, no. 2, pp. 229–237, Apr. 2001.
[63] B. Lindemann, "Taste reception," *Physiol. Rev.*, vol. 76, no. 3, pp. 719–766, Jul. 1996.
[64] P. Besnard, P. Passilly-Degrace, and N. A. Khan, "Taste of fat: A sixth taste modality?" *Physiol. Rev.*, vol. 96, no. 1, pp. 151–176, Jan. 2016.
[65] Y. Liu, Q. Luo, H. Shen, S. Zhuang, C. Xu, Y. Dong, Y. Sun, S. Wang, and H. Deng, "Social media big data-based research on the influencing factors of insomnia and spatiotemporal evolution," *IEEE Access*, vol. 8, pp. 41516–41529, 2020.
[66] X. Cai, H. Zhu, and S. Situ, "Guangdong cooking culture landscapes and their regional differences," *Tropical Geography*, vol. 31, no. 3, pp. 321–327, 2011.

[67] J. Zhu, Y. Xu, Z. Fang, S.-L. Shaw, and X. Liu, ''Geographic prevalence and mix of regional cuisines in Chinese cities,'' *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 5, p. 183, May 2018.

[68] A. Aggarwal, C. D. Rehm, P. Monsivais, and A. Drewnowski, ''Importance of taste, nutrition, cost and convenience in relation to diet quality: Evidence of nutrition resilience among U.S. Adults using national health and nutrition examination survey (NHANES) 2007–2010,'' *Preventive Med.*, vol. 90, pp. 184–192, Sep. 2016.

[69] C. Nekitsing, M. M. Hetherington, and P. Blundell-Birtill, ''Developing healthy food preferences in preschool children through taste exposure, sensory learning, and nutrition education,'' *Current Obesity Rep.*, vol. 7, no. 1, pp. 60–67, Mar. 2018.

[70] G. Fisher, ''Nutrition labeling reduces valuations of food through multiple health and taste channels,'' *Appetite*, vol. 120, pp. 500–504, Jan. 2018.

[71] P. L. Gutierrez, R. M. Garrido, F. F. Sandoval, A. A. Dericke, F. T. Seguel, L. B. Medina, and M. L. Rodriguez, ''Nutritional status and sweet taste preference in Chilean adults,'' *Archivos Latinoamericanos De Nutricion*, vol. 70, no. 4, pp. 247–254, Dec. 2020.

[72] T. Gemousakakis, A. Kotini, P. Anninos, A. Zissimopoulos, and P. Prassopoulos, ''MEG evaluation of taste by gender difference,'' *J. Integrative Neurosci.*, vol. 10, no. 4, pp. 537–545, Dec. 2011.

[73] A. C. Mosca, J. H. F. Bult, and M. Stieger, ''Effect of spatial distribution of tastants on taste intensity, fluctuation of taste intensity and consumer preference of (semi-)solid food products,'' *Food Qual. Preference*, vol. 28, no. 1, pp. 182–187, Apr. 2013.

[74] E. Chamoun, D. M. Mutch, E. Allen-Vercoe, A. C. Buchholz, A. M. Duncan, L. L. Spriet, J. Haines, and D. W. L. Ma, ''A review of the associations between single nucleotide polymorphisms in taste receptors, eating behaviours, and health,'' *Critical Reviews In Food Science And Nutrition*, vol. 58, no. 2, pp. 194–207, 2016.

[75] J. Chandrashekar, M. A. Hoon, N. J. P. Ryba, and C. S. Zuker, ''The receptors and cells for mammalian taste,'' *Nature*, vol. 444, no. 7117, pp. 288–294, Nov. 2006.

[76] G. Li, M. Chen, and G. Zhang, ''The cell biology of mammalian taste,'' *Chem. Life*, vol. 32, no. 5, pp. 456–459, 2012.

[77] H. Omran and C. Schmalstieg, ''Genetic predisposition to diseases of the airways and the lungs,'' *Pneumologe*, vol. 15, no. 4, pp. 229–235, Jul. 2018.

[78] M. Sharafi, S. Rawal, M. L. Fernandez, T. B. Huedo-Medina, and V. B. Duffy, ''Taste phenotype associates with cardiovascular disease risk factors via diet quality in multivariate modeling,'' *Physiol. Behav.*, vol. 194, pp. 103–112, Oct. 2018.

[79] M. Melis, M. Mastinu, G. Sollai, D. Paduano, F. Chicco, S. Magrì, P. Usai, R. Crnjar, B. J. Tepper, and I. T. Barbarossa, ''Taste changes in patients with inflammatory bowel disease: Associations with PROP phenotypes and polymorphisms in the salivary protein, gustin and CD36 receptor genes,'' *Nutrients*, vol. 12, no. 2, p. 409, Feb. 2020.

[80] S. Lockyer, A. Spiro, and S. Stanner, ''Dietary fibre and the prevention of chronic disease–should health professionals be doing more to raise awareness?'' *Nutrition Bull.*, vol. 41, no. 3, pp. 214–231, Sep. 2016.

[81] S. Imathiu, ''Street vended foods: Potential for improving food and nutrition security or a risk factor for food borne diseases in developing countries?'' *Current Res. Nutrition Food Sci. J.*, vol. 5, no. 2, pp. 55–65, Aug. 2017.

[82] R. E. Davis, A. W. Trickey, P. Abrahamse, I. Kato, K. Ward, and A. M. Morris, ''Association of cumulative social risk and social support with receipt of chemotherapy among patients with advanced colorectal cancer,'' *JAMA Netw. Open*, vol. 4, no. 6, Jun. 2021, Art. no. e2113533.

QINYAO LUO is currently pursuing the Ph.D. degree with the School of Geosciences and Info-Physics, Central South University. Her research interests include healthy big data, smart cities, and big data mining.

YU LIU is currently an Associate Professor with the Institute of Environment and Development, Guangdong Academy of Social Sciences. His research interests include point cloud, geographic health big data, and geographic information mining.

MINYAN BI received the master's degree in cartography and geographic information systems from Wuhan University. She is currently with SFMAP Technology (Shenzhen) Ltd. Her research interests include smart cities, 3D GIS, BIM, and geospatial-temporal big data.

XI KUAI is currently a Lecturer with the School of Architecture and Urban Planning, Shenzhen University. His current research interests include geospatial health big data, geographic information semantics, and BIM.

QIN TIAN received the Ph.D. degree in geographical information systems from Wuhan University, China, in 2021. He is currently an Assistant Researcher with the Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources. His research interests include spatial-temporal big data, spatial modeling, and smart cities.

YUKAI SUN received the M.S. degree from Odessa National Medical University, Ukraine, in 2018. He is currently pursuing the Ph.D. degree with the Max Delbrück Centre for Molecular Medicine and the Experimental and Clinical Research Centre, Charité—Universitäts Medicine Berlin. His research interest includes the treatment and prevention of colorectal cancer (CRC).

SIDA ZHUANG received the B.S. degree in geographical information systems from Wuhan University, China, in 2014, and the joint M.S. degree from ITC, The Netherlands, and the University of Southampton, U.K., in 2016. She is currently pursuing the Ph.D. degree with the Applied Geoinformation Laboratory, Humboldt University of Berlin. Her research interests include health geography, spatial modeling, smart cities, and big data mining.