The Association between Diet and Cancer Risk amongst the Chinese:

Evidence from CLHLS 2014

LianLian Chen

Thesis

April 6th, 20

Abstract:

Introduction：

The purpose of this thesis is to study the association between cancer and diets among an aging Chinese population. Most of the existing researches that studies the association between cancer and diet focuses on developed western countries and white populations, and those studies usually focus on diet factors with certain specific diseases, such as breast cancer. This study will fill in this study gap, and is mainly focused on an aging Chinese population, and will study the relationship between diets and cancer in general.

Method:

The dataset for this study is from the Chinese longitudinal healthy longevity survey which was initiated in China since 1998. The Chinese longitudinal healthy longevity survey was a longitudinal study and contains seven waves in total, but this study will focus on the cross-sectional study that was conducted and collected in the year of 2014. The dietary variables that will be used for this study are staple food, fruits, vegetables, food made from beans, salt-preserved vegetables, sugar, tea, fish, meat, garlic, milk products, nuts such as peanuts and walnuts, mushroom or algae, vitamin products and medical plants.

Results:

The mean age of the participants is 85.32, and there are slightly more female than male participants. For people who prefer wheat as the main staple food, the odds of having cancer is 9.11($p<0.01$, 95%CI 1.23,90.9) times the odds of choosing other staple food. Among people who eat fresh fruit occasionally and quiet often(which was the way it measured in the survey), the chance of getting cancer is 96% less likely than people who don't eat that often. Among people who drink tea once a month or more, the odds of having cancer is 77.3 ($p<0.01$, 95%CI 4.31,1857) times more of people don't drink tea that often. Among people who rarely or never eat fish, the odds of having cancer is 67.1 ($p<0.01$, 95%CI 0.40,2.45) times more of people who eat fish. Among people who eat food made from beans at least once a week, the chance of getting cancer is 97%

less likely than people who don't eat that often($p<0.01$, 95%CI 0.0,0.4). There are no significant odds ratio among people who eat salted vegetables. Among people who eat sugar once a month, the chance of getting cancer is 77% less likely than people who don't eat that often($p<0.01$, 95%CI 0.07,0.72), and among people who rarely or never eat sugar, the chance of getting cancer is 61% less likely than people who don't eat that often($p<0.01$, 95%CI 0.12,0.08. Among people who rarely or never eat garlic, the odds of having cancer is 10.01 ($p<0.01$, 95%CI 1.73,192) times of people who eat garlic. The other variables does not show significant odds ratios.

Discussion:

Consumption of fruits, fish, food made from beans and garlic will reduce the chance of getting cancer, which are consistent with findings from previously studies. Other findings such as the association between cancer and vegetable, wheat, tea, sugar, meat and salt preserved vegetable are not consistent with previously studies. Therefore, further analysis and research need to be performed. There are also some data and methodology limitations for this study, such as the population average age is about 84 years old, therefore this could make some findings different from other previous studies that have younger populations.

Introduction:

Cancer is one of the most lethal diseases across the world, and an enormous barrier to increase life expectancy in every country of the world. According to the statistics from the World Health Organization, cancer is the first or second leading cause of death before the age of 70 years in 112 of 183 countries and ranks third or fourth in a further 23 countries.[1] Overall, the burden of cancer incidence and mortality is growing rapidly around the world. However, the distribution of cancer cases and deaths varies by world region. According to the statistics from GLOBOCAN, 58.3% of cancer deaths occur in Asia in 2020, while China accounted for 24% of newly diagnosed cases and 30% of cancer related death worldwide .[2] The current population of Asia is around

4.69 billion in the year of 2021, and is about 59.76% of the world's population. Asian Americans are also the fastest growing population in the US with a 72% growth rate between 2000 and 2015. Therefore, it is very important to study the health issues among this population.

Firstly, it is urgent for countries with a large Asian population, such as US, to provide better healthcare targeting those people. The current healthcare system in western countries such as US does not pay enough attention to Asian health. Some Asian patients in the US reported that they have encountered significant barriers when discussing the use of non-western medical practices with their providers, and they expressed the need for help in navigating the western health care system and obtaining support services.[3] According to some investigations, the reason behind this situation is due to the long history of anti-Asian discrimination from 1882 to 1965[4]. Also, the model minority terms that appeared around 1996 may unconsciously be aggregating Asian American with white American in population studies. Therefore, Asian Americans may be assumed to have similar disease risk profiles to the white population. By accepting the advantages of this model minority label, Asian Americans may also lose the chance to established their own physical and mental health concerns and needs. Secondly, some countries in Asia, such as China, is experiencing transitions in the stages of cancer, meaning the incidence and mortality profile of China are changing from those of developing countries to those of developed countries[5]. Since China has the largest population of the world, any minimum increases in cancer cases and mortality rates indicate significant burden to individuals. Therefore, an understanding of this transition and an analysis of cancer burden are urgently needed in those countries.

There are many articles that emphasize the fact that Asians are not free from cancer burden even though Asian communities have the lowest mortality rate among all racial ethnic groups. In fact, cancer is considered to be the leading cause of death for Asian populations since 2000, while cardiovascular disease being the leading cause of mortality among non-Hispanic whites, non-Hispanic blacks, and Hispanic[6]. There is no question that cancer is the biggest concern to Asian Americans as it is the leading cause of death and bring a lot of burdens to people. The cancer burden among Asian

Americans is unique and varying among different subgroups. The environment, social economic status and life style are very different across different subgroups of Asian communities. Therefore, the prevalence of cancers will be slightly different between them. Lung cancer, nasopharyngeal cancer, liver cancer, gastric cancer, breast cancer is considered to be the most frequent cancers among Asians.[7] Specifically, Lung cancer mortality rate is very high in Chinese Americans, both men and women, according to an Asian Americans cancer mortality rate survey in New York[8]; The nasopharyngeal cancer mortality rate in Chinese men is 15 times that of non-Hispanic white, and is about 12 times as high as non-Hispanic white women[9]; Korean men experienced about 5 times higher incidence rate of getting gastric cancer; the incidence rate in Japanese men was about 4 times higher than white men, and in Japanese women was about 6 times higher[10]. The reason behind this phenomenon is complex, including genetic predisposition, socioeconomic and cultural circumstances. However, dietary could be one of the most important reasons that lead to this high cancer rate.

There are multiple previous researches studied the association between cancer and diet, and it has been estimated that about 30-40 percent of all cancers can be prevented by lifestyle and dietary measures alone[11]. For instance, risk of cancer in relation to dietary pattern, identified using factor analysis, has been investigated in in studies conducted in western populations in the US and Europe, and some Asian countries such as Japan and China. Some studies explored this association between gastric cancer and diet in men and women in the US. According to the result, men who frequently consume high beneficial plant food such as citrus fruit, vegetables and whole grains have lower risk of getting stomach cancer[12]. In women, there is very little association between plant food groups and stomach cancer risk except for increased risk in high fiber vegetables consumption and raw vegetable consumption. In addition to that consuming brown rice and liver would increase the mortality rate of gastric cancer. There are several studies investigated the association between diet and risk of getting breast cancer, and many suggest that diet habits would affect the incidence of breast cancer. According to one article of the related topic, alcohol consumption has a positive relationship with the risk of breast cancer; intake of phytoestrogen may reduce risk of breast cancer; vitamin D can reduce the risk of breast cancer, and lower intake of folate

may be linked to a higher risk of breast cancer.[13] There are also some studies explored the association between Mediterranean dietary, and concluded that a Mediterranean diet appears to be favorable for reduce cancer risk. According to the findings, that up to 25% of colorectal cancers, 15% of breast cancers, and 10% of prostate, pancreatic, and endometrial cancers could be prevented by shifting to a Mediterranean diet[14]. Specifically, increasing consumption of vegetable, fruit, fish and whole grains will decrease the risk of getting cancer, while olive oil, and monounsaturated or unsaturated fats were inversely related to cancer risk. Those findings all shown that there is an association between cancer and dietary pattern, and shifting to a different diet may even change the likelihood of getting cancer.

However, the existing literature findings have a lot of literature gaps. First of all, even though current studies showing that there is a relationship between diet and cancer, the specific association between dietary factors and cancer are sometimes unclear or even inconsistent. For example, though most studies agree that consumption increases the risk of getting breast cancer, but findings for other dietary variables remain uncertain, such as there are conflicting results regarding the association between dietary fat, dairy products and phytoestrogens and breast cancer.

Moreover, despite the extensive studies on the nexus of diet and cancer risk, most of those were conducted developed countries, and generally focused on western white populations, and the relationship between cancer and diet in specific subgroups, such as Asians are largely ignored. Even though there are some researchers found that properly use of traditional Asian nutritional products can assist treatments and recovery, and is essential to cancer healing[15]; and there is a positive relationship between some traditional Korean and Chinese food such as Kimchi[16], there are no study examines carefully about the dietary of Asian communities, and how that could possibly leads to the high cancer rate among them. Different eating habits in different culture may influence the development of cancer differently, especially for Asians that have unique dietary habit. Therefore it is critical to study the relationship between the high cancer rate among Asian population, and their dietary preference.

In addition to that, most of the research on nutrition and cancer has been reductionist; which means, a particular food or a nutrient has been studied in relation to its impact on tumor formation or some other end point of cancer at a particular site in the body. These studies are very helpful in seeing the details of the mechanisms of a specific cancer. However, they do not help give an overall picture in terms of how to prevent cancer on a dietary level. Even less, they

tell little of how to eat when a person already has a cancer and would like to eat a diet that is favorable to their recovery. Therefore, it is important to study a specific dietary pattern that consist of a wide variety of food groups that different from the previous studies. The traditional Chinese diet has the following common characteristics: high consumption of rice, vegetables, bread and other cereals, potatoes, beans, nuts and seeds; low to moderate consumption of dairy products, fish, eggs and poultry; and infrequent consumption of red meat, and is marked by a high intake of rice, vegetables, fruit, and legumes and a low intake of fat and meat[17]. If the dietary components of a Chinese diet have an etiologic role, it will help provide the new empirical evidence for the demonstration of risk relations with diet that is different from the Western.

Thus, the proposed study is to investigate whether dietary intake affects the cancer risk of Chinese, with the focus on staples (such as rice, corn, wheat), fruits and vegetables, meat, eggs, beans (tofu), tea and salt-preserved vegetables. The general research focus of this study is to examine "How some aspects of diet affect the overall cancer rate in a Chinese population"? More specifically the proposed study develops the following hypotheses:

"How does the intake of low fiber affect the cancer prevalence in China?"

"How does the intake of salt-preserved food affect the cancer prevalence in China?"

"How does the intake of food contains high omega 3 affect the cancer prevalence in China?"

The rest of the paper will introduce the dataset of the paper and present the result. The significance of the result and limitations will also be explored in discussion section.

Method:

Data Source and Study Design:

The Chinese longitudinal healthy longevity survey has been initiated in China since 1998. The purpose of this study is to draw social and governmental attention to the oldest population of the society. According to the background, the number of persons aged 80 and older in China is expected to increase to 114 million in 2030.[18]The proportion of old persons aged 65 and older will increase to 16 percent in 2030, and 23 percent in 2050. Additionally, for the time period of 2002-2014, the younger elderly population of the same provinces were interviewed. The baseline survey and the follow-up survey was conducted in half of the cities in 22 of 31 Chinese province, and those cities were randomly selected. The survey constitutes about 85 percent of the total population in China. The baseline survey was originally conducted in 1998, and followed up in 2000,2002, 2005, 2008, 2011, 2012 and 2014.

The Chinese longitudinal healthy longevity survey focused on the oldest population, and provides information on the health, socioeconomic, characteristics, family, lifestyle and demographic profile of this population.

It has conducted seven waves of in-depth surveys between 1998-2014 in 22 provinces in China, using internationally compatible questionnaires. In total, they have conducted 96,805 face- to-face interviews, with 16,557 centenarians, 23,081 nonagenarians, 25,842 octogenarians, 19,650 younger elders aged 65-79 and 11,675 adults aged 35-64. There were data on mortality and length of disability before death for 26,701 participants. For those who died between waves, their data were collected in interviews with a close family member of the deceased

Data collected:

The data was collected through an interview process at interviewee's home, and was collected through face to face interviews that was adapted to the Chinese culture and social context. The data included chronic disease, diet, smoking, alcohol drinking, medical care, nutrition and health-related condition. Participants were asked about their life styles including diets, nutrition, drinking and smoking habits. The survey also included participants' health condition, such as if they suffered from diabetes, heart disease, stroke and cancer. For this study, the dataset that will be used would mostly focus on the dietary, nutrition, cancer rate and other health-related condition.

The wave of the survey used in this study:

The Chinese longitudinal healthy longevity survey was a longitudinal study overall and contains seven waves in total, however, for this study, the dataset that will be used was the cross-sectional study that was conducted and collected in the year of 2014. The reason that this study chooses cross sectional study over longitudinal study is that, it allows researchers to compare many different variables at the same time. Cross sectional study is a commonly used methodology in epidemiology to examine the relationship between disease and exposures. This methodology is able to provide a snapshot of the frequency of a disease in a population at a given time, and is often used to estimate the prevalence and the burden of the disease within a population. One of the advantage of cross sectional study is that it is good for descriptive analyses and for generating hypotheses.[19]

Measures:

The total sample size for this study is 7192. There are three types of variables. The outcome variable was the occurrence of cancer. This variable is a binary variable measured in 0 and 1. In this dataset, it has 4 levels, which are 1= yes, 2= no, 3=do not know and 4 = missing. The study has categorized the variable again and make yes equal to clinical diagnosis of cancer; and other answers equal to no diagnosis of cancer for this variable. The explanatory variables are diet related variables including staple food, fruits, vegetables, food made from beans, salt-preserved vegetables, sugar, tea, fish, meat, garlic, milk products, nuts such as peanuts and walnuts, mushroom or algae, vitamin products and medical plants. For those diet related variable, we are using the result from the kind of food you eat around the age of 60 instead of present, in order to deduce whether eating certain food may lead to cancer, since people may change their eating habit after being diagnosed with cancer. The diet related variables are all categorical variables. For staple food, it was measured in 5 different levels, which are rice, corn(maize), wheat(noodles and bread), rice and wheat, and other. Other diets explanatory variables are all measure in 7 categories, which are rarely or never, don't know, almost every day, not every day, but at least once per week, not every week, but at least once per month, not every month, but occasionally and missing. The control variables are demographics and health behaviors related data including age, gender, marriage status, education level, smoking and drinking habit. Age and education are continuous variables and others are categorical. Education was measured through years of attending school, and the study recoded it into 3 categories: less than high school, high school and more than high school. There are 2 levels for gender which is male and female. Marriage status was measured in 5 levels, currently married and living with spouse; married but not living with spouse; divorced; widowed, and never married. This study has recoded marriage status and made it into 3 categories, which are unmarried, married, and Separated. Smoking and drinking are binary variables. Smoking was measured with either smoking or non-smoking, and drinking was measured with either drinking or non-drinking. All study variables were self-reported. The details of the variables can be accessed the Appendix.

Statistical analysis:

The software that will be used in this study to do statistical analysis will be R. The weights of survey are already conducted by the researchers for producing correct estimates of the average of the entire population.

There will be three types of statistical analysis:

Descriptive statistics will be used to summarize the basic features of the dataset. For continuous variables, it will display mean, standard deviation, median, interquartile range, minimum and maximum. For each categorical variable, the proposed study will show the frequency and percentages by cancer status: diagnosed or not. This helps gain the overview of the participant characteristics.

Correlational analysis will be used to analysis the association between cancer and dietary variables. This study will use bivariate analysis to display the correlation. It will show the frequency of diagnosed of cancer or not for each dietary variable. Variables fruits, vegetables, food made from beans, salt-preserved vegetables, sugar, tea, fish, meat, garlic, milk products, nuts such as peanuts and walnuts, mushroom or algae, vitamin products and medical plants were recoded into 2 levels. Participants' answers that were rarely or never, don't know, missing would collapsed into no, and participants' answers that were almost every day, often, occasionally would collapsed into yes.

logistic regression will be used to estimate odds ratios (ORs) p values, and 95% confidence intervals (CIs) of cancer risk. Logistic regression is a statistical analysis method that is used to obtain odds ratio in the presence of more than one explanatory variable. This method would show the impact of each variable on the odds ratio of the observed event of interest. One of the biggest advantages of logistic regression is that it could avoid confounding effects by analyzing the association of all variables together[20]. For this study, logistic regression will be used since the dependent variable, cancer, of the proposed research is a binary variable. Models will be minimally adjusted for control variables, diet variables and a fully adjusted models include the diet variables and the control variables. This will help generate a better picture of whether diet affects the risk of cancer. The stepwise model selection, and AIC will be use to choose the best model.

Results:

Table 1 displays the basic descriptive statistics of the demographic variables age, gender, marital status, educational level, smoking and drinking. According to the

statistics, the mean age of the participants is 85.32.The minimum age is 47, while the maximum age is 117 years old. There are slightly more female(53.86%) than male (46.11%) participants. Majority of the participants are either married (40.17%) or widowed(58.79%). Only 1% of the participants were people that never married. The majority of the participants have education level less than high school(97.6%), and only one percent of the participants have educational level higher than high school level. 76.08% of the participants reported the habit of smoking, and 88.54% of the participants reported that they drink alcohol regularly.

Table2 shows the bivariate analysis between the outcome variable cancer, and the diets variables fruits, vegetables, food made from beans, salt-preserved vegetables, sugar, tea, fish, meat, garlic, milk products, nuts such as peanuts and walnuts, mushroom or algae, vitamin products and medical plants. The result of this analysis will display the diets variables into 2 categories corresponded to the outcome variable. It shown the percentage and frequency of people who have and don't have cancer with different eating habits.

Table3 shows the result of the logistic regression analysis for 3 models. The first model only contains outcome variable cancer and control variables, and the only variable that is significance in this model is educational level. Among participants who went to high school, the odds of having cancer is 4.9($p<0.001$, 95%CI 1.64,12.0) times more of the odds in people who had educational level less than high school. For model 2, which is the model contains outcome variable cancer and diet variables, the logistic regression shows that among people who prefer wheat as the main staple food, the odds of having cancer is 9.11($p<0.01$, 95%CI 1.23,90.9) times more of the odds of choosing other staple food. The result of Model 3, which is the full model contains outcome variable cancer, control variables and diet variables, however, is not consistent with this result. According to model 3, people who prefer other food (exclude rice, corn, and wheat) as the main stable food have a higher odds 47.7($p<0.01$, 95%CI 5.12,368) of getting cancer than people who choose other 3 type of staple food. According to model 2, among people who rarely or never eat fish, the odds of having cancer is 67.1 ($p<0.01$, 95%CI 0.40,2.45) times more of people who eat fish. According to model 3, among people who eat fresh fruit occasionally and quiet often, the chance of getting cancer is

96% less likely than people who don't eat that often($p<0.01$,95%CI 0.0,0.4). This result is consistent with model 3, which show the odds of getting cancer would be 83%($p<0.01$,95%CI 0.02,0.71) less likely for people who eat fruits quite often. According to model 2, among people who eat meat not every month but occasionally , the chance of getting cancer is 99% less likely than people who eat them often or never eat them($p<0.01$, 95%CI 0.16,2.88). According to model 2,among people who drink tea once a month, the odds of having cancer is 77.3 ($p=0.003$, 95%CI 4.31,1857) times more than people don't drink that often. According to model 2,among people who eat food made from beans at least once a week, the chance of getting cancer is 97% less likely than people who don't eat that often($p<0.01$, 95%CI 0.0,0.4). There are no significant odds ratio among people who eat salted vegetables. According to model 2, among people who eat sugar once a month, the chance of getting cancer is 77% less likely than people who don't eat that often($p<0.01$, 95%CI 0.07,0.72), and among people who rarely or never eat sugar, the chance of getting cancer is 61% less likely than people who don't eat that often($p<0.01$, 95%CI 0.12,0.08). The result from Model 3, which is the full model, also shows a similar relationship as model 2 of the variable sugar. According to model 3, people who eat sugar not every month, but occasionally are 71% less likely than people who don't eat that often($p<0.01$,95%CI 0.08,0.61). In model 3, which is the full model, the result shows that among people who rarely or never eat garlic, the odds of having cancer is 10.01 ($p<0.01$, 95%CI 1.73,192) times of people who eat garlic.

Discussion:

Based on the results of the models, this study finds that certain diet variables has a significant association with the occurrence of cancer. A lot of the results from this study are consistent with other previous studies. Both model 2 and model 3 suggests that those who eat fruits quite often and occasionally will decrease the chance of getting cancer. This finding is consistent with some other studies such as an Italian and Swiss case control studies and a study called "Fruit and vegetable and cancer risk". [22-23] According to the findings from several case control studies, there is about 50% reduction in the risk of getting cancers in people with high intakes of fruits and vegetables. However, our study only shows a positive association between cancer and

fruits, and there is no significant odds ratio for vegetables and cancer. Therefore, the finding only proves part of the hypothesis which is higher fiber is associate with low rates of cancer, since some fruits are high in fiber. According to model 2, people who don't eat fish have a higher chance of getting cancer. Therefore, it answers the hypothesis that intake of food contains high omega 3 could reduce the risk of getting cancer. This finding is consistent with findings from previous studies[24]. One study on the association between fish and cancer suggests that consumption small amount of fish would reduce the risk of getting cancer.[25] Our study also shows that people who eat food made from beans at least once a week have a lower chance of getting cancer compare to people who don't eat that often. This finding is consistent with a previous study which also states that bean consumption is associated with reduced cancer risk in human populations.[26] Our study also finds out that for people who rarely or never eat garlic, the odds of getting cancer is greater than people who eat them. This result is consistent with findings from a previous study, which suggest that intakes of garlic are inversely associated with cancer. [27]

There are some results from our study that are not consistent with previous studies. According to model 2 of our study, for people who eat meat not every month but occasionally, the chance of getting cancer is lower than people who never eat them. However, this finding is not consistent with the results of previous studies. According to one study, processed meat consumption was significantly associated with a 6% greater breast cancer risk, an 18% greater colorectal cancer risk, a 21% greater colon cancer risk, a 22% greater rectal cancer risk, and a 12% greater lung cancer risk.[28] Both model 2 and model 3 also suggests that consumption of sugar regularly will decrease the chance of getting cancer. However, this finding is also inconsistent with the findings from other papers. This inconsistent is probably due to the fact of our dataset mostly focus on the older population of China. According to a paper from pubmed[29] , in 8 of 15 studies on sugary foods and beverages, and a 23-200% higher cancer risk was observed with higher sugary beverage consumption, but since this study is focusing on an older Chinese population, the eating habit of senior Chinese people need to be taken as consideration. Though increasing rapidly recent years, the consumption of fast food is still less common in China compare with western countries,

especially in rural area[30]. Therefore, older people may prefer to stay with traditional diets rather than the change to fast foods as seen in younger groups. We could deduce that the type of meat and sugar that was consumed by this survey participants may be a lot healthier than sugar that come from fast food or processed food. Also, there is a selection bias of our data, for it contains more senior participants. Therefore, a lot of participants with more severe cancers may have already died because of obesity and metabolic syndromes which are also related to sugar. The result from model 2 suggests that eat wheat as the main stable food would increase the chance of getting cancer, and this finding is also inconsistent with the findings from previous studies. According to one study[31] ,whole grain intake is associated with lower risk of total and site-specific cancer, and it has suggested to increase whole grain consumption. However, that article also pointed out the relationship between refined grain intake and cancer risk is not clear. Refined grain food such as noodles and dim sum are very popular in china[32], and may be one explanation behind this observation. However, further study is necessary to be performed on this specific question. The relationship between tea consumption and cancer is not consistent in previous studies, while some studies have shown a protective effect of tea consumption against certain types of cancers, other studies have indicated an opposite effect.[33] According to the result of our study, people who drink tea at least once a month have a higher chance of getting cancer compare to people who don't drink tea. Some previous studies shown that gastric cancer is highly correlated with salt-preserved food,[34] According to one research, the consumption of salted meat, pickled and preserved vegetables are positively associated with gastric cancer, and reduce the consumption of salt and salt processed food is necessary to preventing gastric cancer.[35] Therefore, our study is expecting to show the similar association between salt-preserved vegetable and cancer. However, this association is not being observed in the current study. It may be due to the limitation of the dataset, since most of the participants of this survey are older people. As stated before, they may prefer to stay with traditional diets rather than the change to fast foods as seen in younger groups. Salted-preserved vegetable that was made in a traditional way may be healthier than those manufacturers that we buy at supermarkets nowadays. However, further study is necessary to be performed on this specific question.

There are some limitations of this study. The main limitations would be potential data limitation and methodological limitation. For data limitation, there are some incomplete and interrupted follow-up of individuals, and loss to follow-up over time; Also, the dietary data collected through face-to-face questionnaire interview may not exactly reflect the real diet due to recall bias. There are also some selection bias in this dataset, for most of the participants are older people. Some cancers such as lung cancer where people would likely die before age 80, therefore it would be hard to see whether some associations happened is due to participants' genetics or diets. Another big data limitation is that this dataset lump all cancers together in one outcome variable—whether or not having cancer, therefore it would be hard to interpret some associations, for different cancer has different diagnosis, and it is probably another reason that lead to the inconsistent of the results of our study with some previous studies.

For methodological limitation, Since the outcome and exposure variables are measured at the same time, it is very difficult to establish causal relationships from a cross-sectional study. Another limitation is that, since logistic regression provides odds ratio for each predictor. The odds differ from the risk, and while the odds may appear to be high, the absolute risk may be low.

Table 1 Characteristics of demographic to the Chinese longitudinal healthy longevity survey, From 2014 (n= 7192)

	N (%)[1]
Age (years)[2]	M= 85.32 SD=10.74 Med=85.00 IQR= [77;93] Range=47—117
Gender	
Male	3316 (46.11)
Female	3876 (53.86)
Marital status	
Married	2807 (40.17)
Widowed	4108 (58.79)

Never married	71(0.01)
Education level	
Less high school	7025 (97.6)
High school	69 (0.01)
More than high school	20 (0.01)
Smoking	
yes	1175 (76.08)
no	5858 (23.92)
Drinking	
yes	533(88.54)
no	69 (11.46)

1 Unless otherwise indicated
2 Non-normally distributed
Acronyms: M (mean); Med (median); IQR (interquartile range); SD (standard deviation)

Table 2 Bivariate analysis of diet variables of the Chinese longitudinal healthy longevity survey, From 2014 (n= 7192)

	cancer	
	Yes N%	No N%
Fruit		
Yes	33 (0.05)	5049(0.75)
No	14 (0.02)	1670(0.25)
Sugar		
Yes	31(0.00)	4350(50.33)
No	16(0.00)	2319(14.43)
Vegetable		

Yes	42(0.01)	6461（95.51）
No	5(0.00)	257 (37.99)
Meat		
Yes	40(0.01)	6(0.00)
No	6220(92.21)	479(7.11)
Eggs		
Yes	43(0.01)	6075(90.10)
No	3(0.00)	625(9.26)
Food made from beans		
Yes	40(0.01)	5942(88.12)
No	6(0.00)	755 (11.20)
Salt preserved food		
Yes	20(0.00)	4249(63.04)
No	26(0.00)	2445(36.27)
Tea		
Yes	17(0.00)	2297(34.01)
No	16 (0.00)	4397(65.24)
Garlic		
Yes	30(0.00)	4175(69.93)
No	16 (0.00)	1981(29.38)
Fish		
Yes	38(0.01)	5202(77.42)
No	9(0.00)	1470(21.88)
Milk		
Yes	22(0.00)	2164(32.26)
No	25(0.00)	4502 (67.04)
nuts		
Yes	24(0.00)	3506(52.21)
No	23(0.00)	3162(47.09)
mushroom		
Yes	21(0.00)	3078(45.84)
No	26(0.00)	3590(53.46)
Vitamins		
Yes	6(0.00)	727(10.82)
No	41(0.01)	5940(88.47)
Medical plants		

Yes	5(0.00)	675(10.05)
No	42(0.01)	5992(89.24)

Table 3 Logistic analysis of control, diets and full model predicting cancer occurrence. Chinese longitudinal healthy longevity survey, From 2014 (n= 7192)

Explanatory variables	Model1			Model2			Model3		
	OR*1*	95% CI*1*	p-value	OR*1*	95% CI*1*	p-value	OR*1*	95% CI*1*	p-value
Drinking									
no	—	—							
yes	0.64	0.21, 1.56	0.4						
Smoking									
no	—	—							
yes	1.12	0.45, 2.48	0.8						
Age	1.02	0.99, 1.05	0.2				1.02	0.99, 1.05	0.10
Education									
Less than high school	—	—					—	—	
High school	4.91	1.64, 12.0	0.001*				4.88	1.63, 11.9	0.001
More than high school	0.00		>0.9				0.00		>0.9
Gender									
male	—	—							
female	0.87	0.45, 1.69	0.7						
Stable food									
rice				—	—		—	—	
corn(maize)				1.03	0.06, 5.24	>0.9	3.55	0.18, 24.2	0.3
wheat (noodles and bread etc.)				9.11	1.23, 90.9	0.038*	3.04	0.81, 11.4	0.089

rice and wheat	0.96	0.34, 2.32	>0.9	2.57	0.61, 10.3	>0.9
other	5.50	0.73, 25.7	0.051	47.7	5.12, 368	0.001*
Fish						
almost everyday	—	—				
quite often	0.76	0.31, 1.91	0.6			
occasionally	0.58	0.22, 1.51	0.3			
rarely or never	67.1	0.40, 2.45	0.017*			
vegetables						
almost everyday	—	—				
except winter	0.67	0.29, 1.44	0.3			
occasionally	0.94	0.21, 2.86	>0.9			
rarely or never	2.11	0.53, 6.52	0.2			
Fruit						
almost everyday	—	—		—	—	
quite often	0.11	0.01, 1.12	0.093	0.17	0.02, 0.71	0.019*
occasionally	0.04	0.00, 0.40	0.018*	0.07	0.00, 0.42	0.015*
rarely or never	0.16	0.53, 6.52	0.073	0.53	0.16, 1.70	0.3
Meat						
almost everyday	—	—		—	—	
not every day, but at least once per week	2.09	0.80, 6.52	0.2	2.02	0.80, 6.16	0.2
not every week, but at least once per month	2.05	0.71, 6.83	0.2	2.07	0.75, 6.62	0.2
not every month, but occasionally	0.01	0.16, 2.88	0.016*	0.77	0.19, 2.99	0.7
rarely or never	0.89	0.21, 3.66	0.9	0.57	0.11, 2.42	0.5
Tea						

almost everyday	—	—		—	—	
not every day, but at least once per week	13.7	1.16, 306	0.067	2.14	0.80, 6.16	0.2
not every week, but at least once per month	77.3	4.31, 1,857	0.003*	2.58	0.75, 6.62	0.2
not every month, but occasionally	2.29	0.03, 126	0.7	0.00	0.19, 19.5	0.7
rarely or never	22.8	1.02, 998	0.07	0.38	0.11, 2.42	0.5
Food made from beans						
almost everyday	—	—				
not every day, but at least once per week	0.03	0.00, 0.40	0.01*			
not every week, but at least once per month	0.03	0.00, 0.51	0.02*			
not every month, but occasionally	0.09	0.03, 1.31	0.08			
rarely or never	0.36	0.02, 5.32	0.5			
Salted Vegetable						
almost everyday	—	—				
not every day, but at least once per week	0.49	0.05, 3.58	0.5			
not every week, but at least once per month	0.03	0.00, 0.79	0.081			
not every month, but occasionally	0.74	0.05, 7.85	0.8			
rarely or never	0.10	0.01, 0.95	0.057			
Sugar						
almost everyday	—	—		—	—	
not every day, but at least once per week	0.39	0.14, 1.09	0.070	0.36	0.13, 0.99	0.048
not every week, but at least once per month	0.24	0.07, 0.72	0.012*	0.29	0.09, 0.81	0.020
not every month, but occasionally	0.23	0.08, 0.65	0.006	0.22	0.08, 0.61	0.004*
rarely or never	0.31	0.12, 0.80	0.013*	0.41	0.18, 0.99	0.039
Garlic						
almost everyday	—	—		—	—	

not every day, but at least once per week	1.77	0.65, 5.34	0.3	1.84	0.17, 40.3	0.6
not every week, but at least once per month	1.90	0.59, 6.32	0.3	2.48	0.22, 55.8	0.5
not every month, but occasionally	1.63	0.52, 5.35	0.4	6.20	0.37, 126	0.11
rarely or never	2.57	0.93, 7.91	0.079	10.1	1.73, 192	0.033*
Nuts						
almost everyday	—	—				
not every day, but at least once per week	1.06	0.24, 7.53	>0.9			
not every week, but at least once per month	0.03	0.00, 1.74	0.070			
not every month, but occasionally	0.73	0.17, 5.22	0.7			
rarely or never	0.87	0.22, 5.92	0.9			
Vitamins						
almost everyday	—	—				
not every day, but at least once per week	1.39	0.05, 37.7	0.8			
not every week, but at least once per month	2.25	0.19, 51.8	0.5			
not every month, but occasionally	0.79	0.07, 18.5	0.9			
rarely or never	1.07	0.19, 20.2	>0.9			
Milk						
almost everyday	—	—				
not every day, but at least once per week	0.98	0.41, 2.31	>0.9			
not every week, but at least once per month	1.03	0.31, 2.91	>0.9			
not every month, but occasionally	0.76	0.23, 2.18	0.6			
rarely or never	0.57	0.24, 1.33	0.2			
Mushroom						

almost everyday	—	—	
not every day, but at least once per week	1.82	0.32, 35.0	0.6
not every week, but at least once per month	1.39	0.23, 27.2	0.8
not every month, but occasionally	0.89	0.15, 17.4	>0.9
rarely or never	0.63	0.11, 12.2	0.7

1 OR = Odds Ratio, CI = Confidence Interval

Reference:

Zeng, Yi, Vaupel, James, Xiao, Zhenyu, Liu, Yuzhi, and Zhang, Chunyuan. Chinese Longitudinal Healthy Longevity Survey (CLHLS), 1998-2014. Inter-university Consortium for Political and Social Research [distributor], 2017-04-11. https://doi.org/10.3886/ICPSR36692.v1

Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. American Cancer Society Journals. https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21660. Published February 4, 2021. Accessed December 18, 2021.

Zhang R, Wang Z, Fei Y, et al. The difference in nutrient intakes between Chinese and Mediterranean, Japanese and American diets. Nutrients. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4488807/. Published June 9, 2015. Accessed December 18, 2021.

Carlo La Vecchia, Association between Mediterranean dietary patterns and cancer risk, *Nutrition Reviews*, Volume 67, Issue suppl_1, 1 May 2009, Pages S126–S129, https://doi.org/10.1111/j.1753-4887.2009.00174.x

Kotepui M. Diet and risk of breast cancer. Contemporary oncology (Poznan, Poland). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4829739/. Published 2016. Accessed December 18, 2021.

Screening Oof C. Changing profiles of cancer burden worldwide and in China:... : Chinese Medical Journal. LWW. https://journals.lww.com/cmj/Fulltext/2021/04050/Changing_profiles_of_cancer_burden_worldwide_and.5.aspx. Accessed December 18, 2021.

Taylor VM, Ko LK, Hwang JH, Sin MK, Inadomi JM. Gastric cancer in asian american populations: a neglected health disparity. *Asian Pac J Cancer Prev*. 2014;15(24):10565-10571. doi:10.7314/apjcp.2014.15.24.10565

Lee E, Liu L, Zhang J, et al. Stomach Cancer Disparity among Korean Americans by Tumor Characteristics: Comparison with Non-Hispanic Whites, Japanese Americans, South Koreans, and Japanese. *Cancer Epidemiol Biomarkers Prev*. 2017;26(4):587-596. doi: 10.1158/1055-9965.EPI-16-0573

Hishida A, Matsuo K, Goto Y, et al. Smoking behavior and risk of Helicobacter pylori infection, gastric atrophy and gastric cancer in Japanese. *Asian Pac J Cancer Prev*. 2010;11(3):669-673.

Woo HDP. Diet and cancer risk in the Korean population: A meta-analysis. Asian Pacific Journal of Cancer Prevention. https://www.koreascience.or.kr/article/JAKO201435053629248.page. Accessed November 10, 2021.

Hong Xu, Houen Xu, 2006 - journals.sagepub.com. https://journals.sagepub.com/doi/10.1177/117863370600100001. Accessed November 10, 2021.

McCullough ML, Robertson AS, Jacobs EJ, Chao A, Calle EE, Thun MJ. A prospective study of diet and stomach cancer mortality in United States men and women. Cancer Epidemiology, Biomarkers & Prevention. https://cebp.aacrjournals.org/content/10/11/1201. Published November 1, 2001. Accessed November 10, 2021.

Kweon S-S, Shu X-O, Xiang Y, et al. One-carbon metabolism dietary factors and distal gastric cancer risk in Chinese women. Cancer Epidemiology, Biomarkers & Prevention. https://cebp.aacrjournals.org/content/23/7/1374. Published July 1, 2014. Accessed November 10, 2021.

Thompson CA, Gomez SL, Hastings KG, et al. The burden of cancer in Asian Americans: A report of national mortality trends by Asian ethnicity. Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5218595/. Published October 2016. Accessed November 10, 2021.

Lee RJ, Madan RA, Kim J, Posadas EM, Yu EY. Disparities in cancer care and the Asian American population. The Oncologist. https://theoncologist.onlinelibrary.wiley.com/doi/10.1002/onco.13748. Published March 20, 2021. Accessed November 10, 2021.

Chen MS. Cancer health disparities among Asian Americans. American Cancer Society Journals. https://acsjournals.onlinelibrary.wiley.com/doi/full/10.1002/cncr.21501. Published November 3, 2005. Accessed November 10, 2021.

Ngo-Metzger Q, Massagli MP, Clarridge BR, et al. Linguistic and cultural barriers to care. Journal of general internal medicine. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1494812/. Published January 2003. Accessed April 4, 2022.

Jin H, Pinheiro PS, Xu J, Amei A. Cancer incidence among Asian American populations in the United States, 2009-2011. International journal of cancer. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5283572/. Published May 1, 2016. Accessed April 4, 2022.

Kotepui M. Diet and risk of breast cancer. Contemporary oncology (Poznan, Poland). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4829739/. Published 2016. Accessed April 4, 2022.

Donaldson MS. Nutrition and cancer: A review of the evidence for an anti-cancer diet - nutrition journal. BioMed Central. https://nutritionj.biomedcentral.com/articles/10.1186/1475-2891-3-19. Published October 20, 2004. Accessed April 4, 2022.

Healthy Longevity in China: Demographic, Socioeconomic and Psychological Dimensions. Lieu de publication inconnu: Springer; 2009.

Health knowledge. https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/cs-as-is. Accessed April 5, 2022.

S; S. Understanding logistic regression analysis. Biochemia medica. https://pubmed.ncbi.nlm.nih.gov/24627710/#:~:text=Logistic%20regression%20is%20used%20to,the%20observed%20event%20of%20interest. Accessed April 6, 2022.

Key TJ. Fruit and vegetables and cancer risk. British journal of cancer. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3039795/. Published January 4, 2011. Accessed May 4, 2022.

Campos-Vega R, Oomah BD, Loarca-Piña G, Vergara-Castañeda HA. Common beans and their non-digestible fraction: Cancer inhibitory activity-an overview. Foods (Basel, Switzerland). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5302293/. Published August 2, 2013. Accessed May 4, 2022.

Miraghajani M, Rafie N, Hajianfar H, Larijani B, Azadbakht L. Aged garlic and cancer: A systematic review. International journal of preventive medicine. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6212616/. Published September 17, 2018. Accessed May 4, 2022.

Fernandez E;Chatenoud L;La Vecchia C;Negri E;Franceschi S; Fish consumption and cancer risk. The American journal of clinical nutrition. https://pubmed.ncbi.nlm.nih.gov/10393143/. Accessed April 6, 2022.

Farvid MS;Sidahmed E;Spence ND;Mante Angua K;Rosner BA;Barnett JB; Consumption of red meat and processed meat and cancer incidence: A systematic review and meta-analysis of prospective studies. European journal of epidemiology. https://pubmed.ncbi.nlm.nih.gov/34455534/#:~:text=Processed%20meat%20consumption%20was%20significantly,12%25%20greater%20lung%20cancer%20risk. Accessed May 4, 2022.

Lin SH;Li YH;Leung K;Huang CY;Wang XR; Salt processed food and gastric cancer in a Chinese population. Asian Pacific journal of cancer prevention: APJCP. https://pubmed.ncbi.nlm.nih.gov/25040991/#:~:text=Our%20results%20suggest%20that%20consumption,measure%20to%20preventing%20gastric%20cancer. Accessed April 6, 2022.

Ji XN;Huang M;Yao SH;Qi JY;Onwuka JU;Wang Y;Wang XM;Chen Y;Wu XY;Liu LY;Wang SH;Zhou M;He YH;Feng RN; Refined grains intake in high fat, high protein, low carbohydrate and low energy levels subgroups and higher likelihood of abdominal obesity in Chinese population. International journal of food sciences and nutrition. https://pubmed.ncbi.nlm.nih.gov/32233701/. Accessed April 6, 2022.

Turati F;Rossi M;Pelucchi C;Levi F;La Vecchia C; Fruit and vegetables and cancer risk: A Review of Southern European Studies. The British journal of nutrition. https://pubmed.ncbi.nlm.nih.gov/26148912/. Accessed April 6, 2022.

Gaesser GA. Whole grains, refined grains, and cancer risk: A systematic review of Meta-analyses of observational studies. Nutrients. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7762239/. Published December 7, 2020. Accessed April 6, 2022.

Lin S;Xu G;Chen Z;Liu X;Li J;Ma L;Wang X; Tea drinking and the risk of esophageal cancer: Focus on tea type and drinking temperature. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP). https://pubmed.ncbi.nlm.nih.gov/32740163/#:~:text=Drinking%20very%20hot%20tea%20(%3E65,)%20relative%20to%20non%2Ddrinkers. Accessed May 4, 2022.