

Research Paper

# Nomograms and risk scores for predicting the risk of oral cancer in different sexes: a large-scale case-control study

Fa Chen<sup>1,2\*</sup>, Lisong Lin<sup>3\*</sup>, Lingjun Yan<sup>1,2</sup>, Fengqiong Liu<sup>1,2</sup>, Yu Qiu<sup>3</sup>, Jing Wang<sup>4</sup>, Zhijian Hu<sup>1,2</sup>, Junfeng Wu<sup>1,2</sup>, Xiaodan Bao<sup>1,2</sup>, Liangkun Lin<sup>1,2</sup>, Rui Wang<sup>1,2</sup>, Guoxi Cai<sup>5</sup>, Kiyoshi Aoyagi<sup>6</sup>, Lin Cai<sup>1</sup>, Baochang He,<sup>1,2✉</sup>

1. Department of Epidemiology and Health Statistics, Fujian Provincial Key Laboratory of Environment Factors and Cancer, School of Public Health, Fujian Medical University, Fuzhou, China;
2. Key Laboratory of Ministry of Education for Gastrointestinal Cancer, Fujian Medical University, Fuzhou, China;
3. Department of Oral and Maxillofacial Surgery, the First Affiliated Hospital of Fujian Medical University, Fuzhou, China;
4. Laboratory Center, School of Public Health, Fujian Medical University, Fuzhou, China;
5. Nagasaki Prefectural Institute of Environmental Research and Public Health; Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan;
6. Department of Public Health, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan.

\*These authors contributed equally to this work.

✉ Corresponding author: Baochang He, Department of Epidemiology and Health Statistic, School of Public Health, Fujian Medical University, 1 Xueyuan Road, Fuzhou 350108, China. Tel: 86-18905004513; Fax: 86-0591-22862510; E-mail: hbc517@163.com

© Ivyspring International Publisher. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2017.12.18; Accepted: 2018.04.29; Published: 2018.06.22

## Abstract

**Background:** Although previous studies have explored the associations of modifiable lifestyle factors with oral cancer risk, few studies integrated these factors and established predictive tools for oral cancer risk in different sexes.

**Methods:** Using a case-control study design, a total of 978 oral cancer cases and 2646 healthy controls were recruited in this study. Nomograms were constructed according to significant factors in multivariable logistic regression. Risk scores were calculated based on the nomograms and quantified the risk of oral cancer using restricted cubic spline.

**Results:** Multivariate analyses demonstrated that smoking, alcohol drinking, tea, intake of fish, seafood, vegetables, fruits, teeth loss, regular dental visits and repetitive dental ulcer were independent factors for male oral cancer. Passive smoking, age at first intercourse, cooking oil fumes exposure, tea, intake of beans, vegetables, fruits, teeth loss, regular dental visits and repetitive dental ulcer were associated with female oral cancer. Then, two nomograms were developed for predicting the probability of oral cancer in men and women with the C-index of 0.768 (95% CI: 0.723-0.813) and 0.700 (95% CI: 0.635-0.765), respectively. Restricted cubic splines graphically revealed the risk of oral cancer in individuals with different risk scores. Moreover, the risk escalated continuously with the increasing number of the risk scores among both sexes.

**Conclusions:** Combining nomograms with risk scores developed in this study could precisely predict oral cancer occurrence and provide an accurate risk assessment.

Key words: Oral cancer; lifestyle factors; nomogram; risk score; sex

## Introduction

Oral cancer, one of the most common malignant tumors in head and neck, is still a major public health burden in developing countries. Although the incidence of oral cancer in men is higher than that of women[1], the ratio of men to women has been gradually decreasing from 6:1 in 1950 to 2:1 in 2015[2],

suggesting that the incidence of female oral cancer showed a more significant upward trend[3]. Given the differences in physiology and lifestyles of both sexes, it would be valuable to separately identify the influencing factors for oral cancer in men and women.

There is accumulating epidemiologic evidence

that modifiable lifestyle factors exert a significant impact on oral cancer risk but the associations vary by sex. Tobacco smoking was a major risk factor for oral cancer, which has been well-established previously[4]. Muscat et al.[5] further indicated that female smokers had a greater risk of oral cancer than male smokers. Many previous studies have pointed to a strong protective effect of tea consumption on oral cancer risk[6, 7]. Our recent study demonstrated that women could derive the more benefit from tea drinking than men[8]. Besides, passive smoking and cooking oil fumes (COF) exposure were reported to significantly increase the risk of oral cancer in Chinese women, while male exposure rates were relatively low[9].

Compared with the single risk factor, integrating several significant factors may be more efficient and precise to evaluate the role of modifiable lifestyle factors in the etiology of oral cancer. Nomogram is a graphical representation tool commonly used to predict the prognosis of patients by combining significant factors[10]. However, so far, few studies have established nomograms applicable to Chinese population for predicting the probability of the occurrence of oral cancer in men and women, respectively. Moreover, continuous quantification of oral cancer risk according to the total score from nomogram has also been rarely reported.

Therefore, in this large-scale case-control study, we comprehensively investigated the gender differences in the modifiable lifestyle factors of oral cancer, and further developed nomograms and risk scores for predicting the risk of oral cancer in different sexes.

## Materials and methods

### Study subjects

From September 2010 to March 2017, a hospital-based case-control study was conducted in Fujian province, China, which has been described previously[11]. In brief, primary oral cancer patients were consecutively recruited from the First Affiliated Hospital of Fujian Medical University with newly diagnosed and histopathology confirmed. Those with recurrent or metastasized oral cancer were excluded. Finally, a total of 978 eligible patients (380 tongue, 135 buccal, 128 gingiva, 72 floor of mouth, 69 palate, 34 lip, and 160 unspecified or overlapping) were included. During the study period, 2646 healthy controls were randomly selected from the health examination center of the same hospital with no any history of cancer or cancer-related diseases. The healthy status of control subjects was determined based on the results of physical examination. Controls

were frequency-matched to cases by age ( $\pm 5$  years). All subjects were Chinese Han population and aged between 20 and 80 years old, and resident in Fujian province. The recruiting rates were 97.3% for cases and 91.6% for controls. Written informed consent was obtained from all participants. This study was approved by the Institutional Review Board (IRB) of Fujian Medical University (Fuzhou, China) and have been performed in line with the ethical standards in the 1964 Declaration of Helsinki and its later amendments.

### Data collection

Data were collected from all subjects through face-to-face interviews using a structured lifestyle questionnaire. Information collected included socio-demographic characteristics; active and passive smoking, alcohol drinking, tea related variables (tea consumption, tea types, tea temperature, and tea concentration, etc.), dietary habits (intake frequency of red meat, domestic meat, fish, seafood (shrimp/crab/shellfish), milk and dairy products, eggs, vegetables, fruits, beans and soy products, pickled food, etc.), COF exposure, oral hygiene status (frequency of tooth brushing, the number of teeth loss, dentures wearing, regular dental visits, repetitive dental ulcer), residential history, family history of cancer.

### Variables definition

Smokers were defined as those who had smoked at least 100 cigarettes during their lifetime. Passive smokers were defined as subjects exposed to environmental tobacco smoke over 15 minutes per day. Alcohol drinking was defined as drinking alcohol at least once per week continuously for at least 6 months. Tea drinkers were defined as having at least 1 cup/week for six months or more. COF exposure was defined as those who cooked food in deep-frying or stir-frying way for more than twice per week with eyes or throat irritation. To assess intensity of COF exposure, participants were interviewed about the fumes in the kitchen when cooking, selecting an answer from three options: "no exposure", "light" or "heavy".

### Statistical analysis

Distribution of baseline characteristic between cases and controls was compared using chi-square test. Some frequencies categories or continuous variables were grouped according to the median or tertiles of the controls. Unconditional logistic regression was used to estimate the odds ratios (ORs) and 95% confidence intervals (CIs) for the associations of modifiable lifestyle factors with oral cancer risk in both sexes. The nomograms for the prediction of the

probability of male and female oral cancer were established with the selected independently significant variables. The discrimination and predictive ability of the nomograms were assessed by Harrell's concordance index (C-index). Calibration curves was analyzed by plotting the nomogram-predicted and the actual probability of the occurrence of oral cancer. Restricted cubic spline was performed to assess the non-linear relationship between the total risk scores and the risk of oral cancer. All statistical analyses were conducted using R (version 3.1.1). Statistical significance was considered at  $P < 0.05$ .

## Results

Of the 3624 of oral cancer patients and controls, 1924 (53.09%) were men, and 1700 (46.91 %) were women. The demographic information of the cases and controls in both sexes are listed in Table 1. Age, education levels and marital status were similarly distributed among cases and controls (all  $P > 0.05$ ). Compared to controls, oral cancer patients were more likely to be rural settings and to have a family history of cancer.

The potential risk factors of oral cancer were shown in Supplementary Table S1, Table S2, and Table S3. Then, multivariate logistic regression was used to select independent factors for oral cancer in men and women. As shown in Table 2, smoking and alcohol drinking were specific risk factors for oral cancer in men but not in women. The protective effects of more consumption of fish ( $\geq 3$  times/week) and seafood ( $\geq 1$  time/week) on oral cancer were also only observed in men. On the contrary, passive smoking, age at first intercourse less than 22 and cooking oil fumes exposure were specific risk factors for female oral cancer. Moreover, more intake of beans ( $\geq 1$  time/week) was particular protective factor for women. Additionally, the similar magnitudes of

the associations related to oral hygiene were observed across genders.

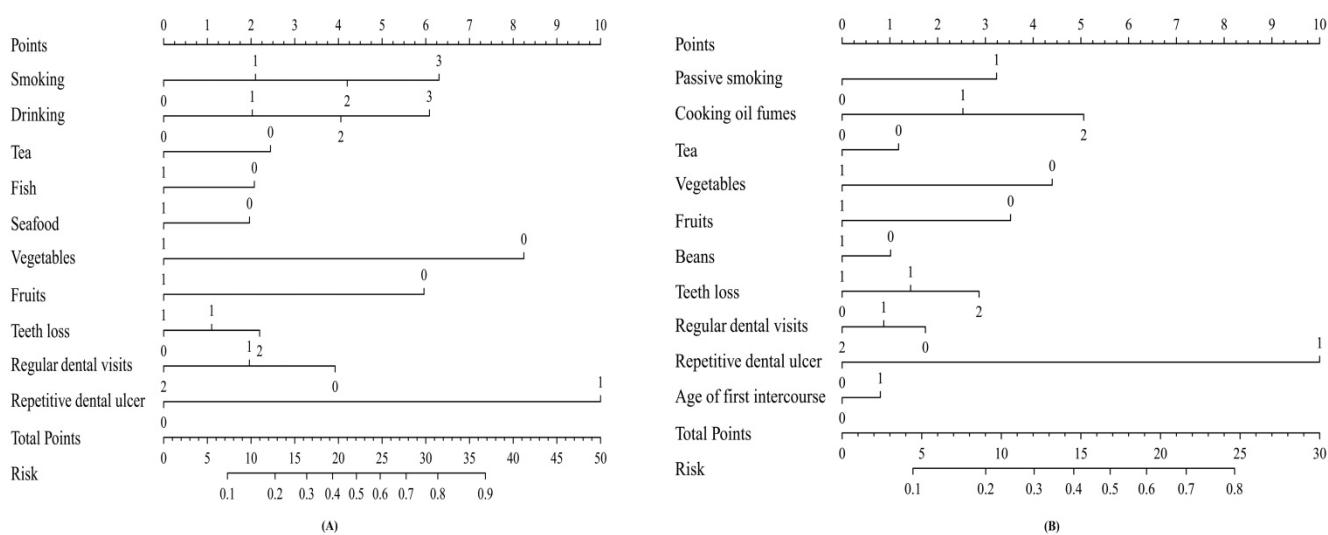
Next, based on these independent factors, we developed two nomograms for predicting the probability of the occurrence of oral cancer in men and women respectively (Figure 1). Each point of independent variables can be determined according to the intersection of the vertical line drawn from the variable to the point axis. Then, the total risk score was calculated by adding each variable point. The probability of the occurrence of oral cancer can be read on the total point axis. The Harrell's C-index of nomogram was 0.768 (95% CI: 0.723-0.813) for men, and 0.700 (95% CI: 0.635-0.765) for women. The calibration curves for men and women all indicated that the predicted probabilities by the nomogram were good match with the actual observation (Supplementary Figure S1, Figure S2).

**Table 1.** Demographic characteristics of oral cancer patients and controls

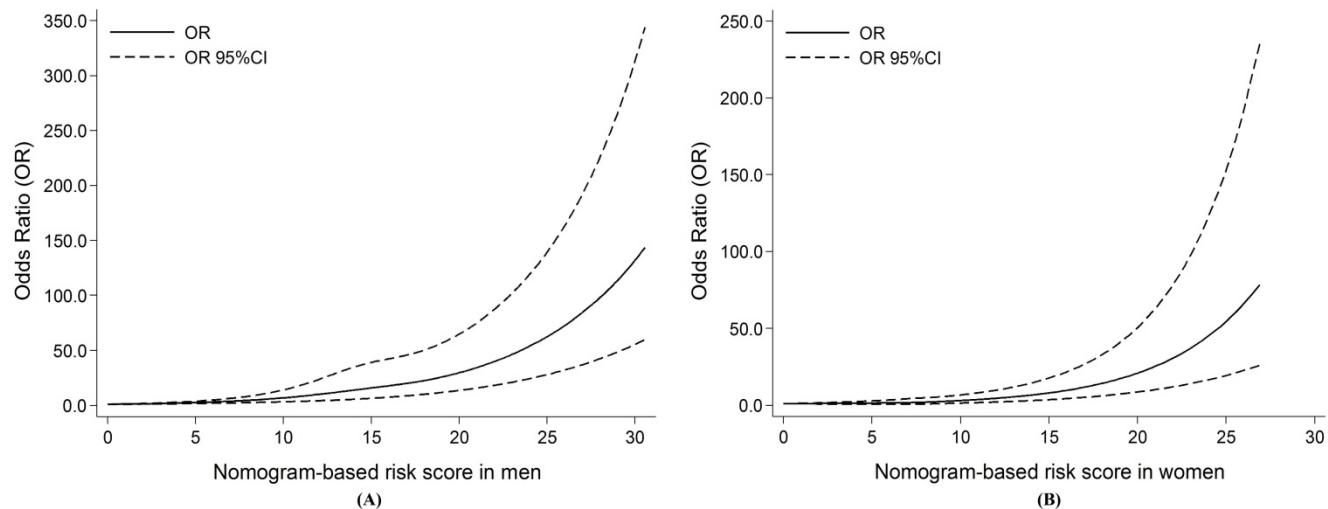
Variable	Men			Women		
	Cases n(%)	Controls n(%)	P	Cases n(%)	Controls n(%)	P
Age(years)			0.184			0.808
<60	351(56.98)	703(53.75)		183(50.55)	686(51.27)	
≥60	265(43.02)	605(46.25)		179(49.45)	652(48.73)	
Education Level			0.126			0.690
Illiteracy	33(5.36)	77(5.89)		77(21.27)	300(22.42)	
Primary and Middle	409(66.40)	806(61.62)		219(60.50)	776(58.00)	
High school and above	174(28.24)	425(32.49)		66(18.23)	262(19.58)	
Marital status			0.070			0.480
Married	568(92.21)	1172(89.60)		320(88.40)	1200(89.69)	
Other	48(7.79)	136(10.40)		42(11.60)	138(10.31)	
Residence			0.033			<0.001
Rural	352(57.14)	814(62.23)		204(56.35)	937(70.03)	
Urban	264(42.86)	494(37.77)		158(43.65)	401(29.97)	
Family history of cancer			<0.001			
No	507(82.31)	1162(88.84)		309(85.36)	1235(92.30)	
Yes	109(17.69)	146(11.16)		53(14.64)	103(7.70)	

**Table 2.** Multivariate logistic regression analysis for the independent risk factors of oral cancer in men and women

Men				Women			
Variable	Categories	Reference	OR(95%CI)	Variable	Categories	Reference	OR(95%CI)
Smoking (pack-year)	<20	Never	1.47(1.04-2.08)	Passive smoking	Yes	No	1.77(1.35-2.32)
	20~		1.62(1.17-2.24)	Cooking oil fume exposure	Middle	Light and below	1.78(1.31-2.41)
	40~		2.84(2.05-3.93)		Heavy		2.60(1.19-5.70)
Alcohol drinking (g/day)	<20	Never	1.05(0.76-1.45)	Tea consumption	Yes	No	0.71(0.51-0.99)
	20~		1.59(1.13-2.23)	Vegetables (times/day)	≥1	<1	0.49(0.30-0.79)
	60~		2.52(1.78-3.57)	Fruits (times/week)	≥3	<3	0.52(0.40-0.69)
Tea consumption	Yes	No	0.66(0.52-0.84)	Beans (times/week)	≥1	<1	0.72(0.54-0.97)
Fish (times/week)	≥3	<3	0.72(0.57-0.92)	Teeth loss	<5	0	1.59(1.07-2.36)
Seafood (times/week)	≥1	<1	0.66(0.51-0.86)		≥5		2.10(1.32-3.35)
Vegetables (times/day)	≥1	<1	0.30(0.21-0.44)	Regular dental visits (times/year)	<1	Never	0.53(0.35-0.81)
Fruits (times/week)	≥3	<3	0.38(0.30-0.49)		≥1		0.45(0.22-0.93)
Teeth loss	<5	0	1.67(1.22-2.28)	Repetitive dental ulcer	Yes	No	6.00(3.67-9.80)
	≥5		1.69(1.19-2.42)	Age of first intercourse	≤22	>22	1.50(1.12-2.01)
Regular dental visits (times/year)	<1	Never	0.50(0.33-0.75)				
	≥1		0.47(0.25-0.89)				
Repetitive dental ulcer	Yes	No	4.76(2.75-8.21)				



**Figure 1.** Nomogram predicting the probability of oral cancer. (A) For men; (B) For women.



**Figure 2.** Restricted cubic splines for the nonlinear relationships of the risk of oral cancer and increasing risk scores. (A) For men; (B) For women. Solid line shows the odds ratios (ORs), and dotted line represents the 95% confidence intervals (CIs).

According to the above method, we further calculated the total risk scores of all subjects. Restricted cubic splines graphically demonstrated that the risk of oral cancer successively escalated with the increasing number of the total risk scores among both sexes (Figure 2). Furthermore, according to the tertiles of controls, the total scores were divided into three risk groups in men (0-10.34, low; 10.35-16.38, moderate; 16.39-44.88, high), and in women (0-6.15, low; 6.16-9.18, moderate; 9.19-27.97, high), respectively. Individuals in moderate and high risk groups had significantly increased risk of oral cancer compared with those in low risk group among both sexes (all  $P < 0.05$ , Table 3).

## Discussion

Overall, this large-scale case-control study is one of the few to comprehensively evaluate the different roles of lifestyle-related factors in oral cancer among

men and women. We successfully established two practical nomograms for predicting the probability of the occurrence of oral cancer in men and women. Two useful risk scores were also developed for continuously quantifying the risk of oral cancer.

**Table 3.** Association between total risk scores and risk of oral cancer in men and women

Variable	Men		Women	
	Cases/Controls	OR(95%CI) <sup>a</sup>	Cases/Controls	OR(95%CI) <sup>a</sup>
<b>Total risk scores</b>				
Low risk	45/433	1.00	53/442	1.00
Moderate risk	124/439	2.97(2.05-4.31)	76/438	1.71(1.17-2.52)
High risk	447/436	11.89(8.39-16.85)	224/440	5.66(3.98-8.05)

<sup>a</sup> Adjustment for age, education, marital status, residence, family history of cancer.

Although previous studies have explored the associations of modifiable lifestyle factors with oral cancer risk, few evaluated the comprehensive effect of these factors and established the predictive model. To

make up for this shortcoming, this study developed two reliable nomograms with high predictive ability according to the conventional obtainable modifiable lifestyle factors. By providing an intuitive line drawings, our nomogram can easily calculate each individual's probability of the occurrence of oral cancer. The nomogram was initially introduced in the field of oncology research in 1998 by Kattan et al.[12]. Subsequently, nomograms were used to predict the probability of various clinical prognostic events (such as recurrence, metastasis, death, etc.) for various kinds of cancer patients[13-15]. Until now, to our knowledge, this study is the first to construct nomogram for predicting the occurrence of oral cancer.

In addition to predicting the probability of oral cancer, several risk models or risk scores have been constructed in previous studies for risk assessment. Liu et al.[16] developed a peaks random forest model according to exfoliative cytology and clinical data for predicting oral cancer risk. Krishna et al.[17] built a risk score model using lifestyle data with a small samples in India. In the present study, we further calculated risk scores to quantitatively predict the risk of oral cancer using restrictive cubic splines. Using the risk score of each patient, oral cancer risk can be visualized on a graphical curve. Compared with our previous proposed environmental exposure index[18], combining nomograms with risk scores would be useful in clinical settings, and could demonstrate an easily accessible visual representation of the predictive probability and the estimated risk of oral cancer. Indeed, they would assist in the screening of individuals in high risk and be helpful to individualized prevention of oral cancer.

Notably, among the significant modifiable lifestyle factors in the predictive models, the favorable roles of fish and seafood consumption were only emerged in men. Fish and seafood are widely consumed in the southeast of China, which are abundant in polyunsaturated fatty acids, and have the effects of anti-inflammatory as well as reducing the production of free radicals and carcinogens[19, 20]. However, little is known about the particular mechanism of fish and seafood on male oral cancer, and more additional research on these associations are warranted.

Conversely, increased consumption of beans was only seen to be protective against oral cancer in women. This finding is in agreement with the results of studies on gastric, colorectal and breast cancer in women[21-23]. Additionally, younger age at first intercourse was found to be specific risk for female oral cancer, which is consistent with a previous study[24]. Anaya-Saavedra et al.[25] indicated that

earlier age at sexual debut could elevate the risk of oral HPV infection. Our recent study also observed an interaction between the first sexual age and oral HPV infection for oral squamous cell carcinoma[26]. Therefore, one possible explanation is that earlier first sexual age may increase the risk of oral cancer by facilitating oral HPV infection.

Several limitations in this study warrant mention. First, since this study is a case-control study, the potential selection bias and recall bias are major concern. To minimize the possibility of the biases, all cases and controls were recruited in the same hospital with a strict criteria and high participation rates. Moreover, all cases were newly diagnosed, and the definitions of all variables were clearly stated during the interview. Second, as a single-center study, it is uncertain whether these predictive tools are applicable to other populations with different cultures and ethnic groups. Although our nomograms were initially validated using bootstraps with 1000 resamples, future work with larger cohorts in multi-center are still needed to externally validate our results. Third, to make our predictive tools widely accepted, we only taken into account several easily obtained lifestyle factors. Hopefully, other known serological and genetic biomarkers would be included in future studies in order to improve the discriminatory ability of nomograms and risk scores.

In conclusion, the present study provided further evidence on the different roles of lifestyle-related factors in the etiology of oral cancer among different sexes. The nomograms and risk scores constructed in this study can provide an accurate prediction of the probability and the estimated risk of oral cancer. We believe that these predictive tools would be useful in conjunction with screening of high-risk population and primary prevention of oral cancer.

## Acknowledgments

This study was funded by the High-level Talents research Start-up Project of Fujian Medical University (No.XRCZX2018001), Scientific Research Talents Training Project of Health and Family Planning Health Commission in Fujian Province (No.2018-1-71, No.2017-ZQN-57), Joint Funds for the Innovation of Science and Technology of Fujian province (No.2017Y9103, No.2016Y9033).

## Supplementary Material

Supplementary figures and tables.

<http://www.jcancer.org/v09p2543s1.pdf>

## Competing Interests

The authors have declared that no competing interest exists.

## References

1. Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. *Oral oncology*. 2009; 45: 309-16.
2. Fédération Dentaire International. The Challenge of Oral Disease- A Call for Global Action. Geneva: FDI World Dental Federation; 2015.
3. Chaturvedi AK, Anderson WF, Lortet-Tieulent J, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. *Journal of clinical oncology*. 2013; 31: 4550-9.
4. Radiot L, Paget-Bailly S, Cyr D, et al. Tobacco smoking, alcohol drinking and risk of oral cavity cancer by subsite: results of a French population-based case-control study, the ICARE study. *European journal of cancer prevention*. 2013; 22: 268-76.
5. Muscat JE, Richie JP, Jr., Thompson S, et al. Gender differences in smoking and risk for oral cancer. *Cancer research*. 1996; 56: 5192-7.
6. Radiot L, Paget-Bailly S, Menvielle G, et al. Tea and coffee consumption and risk of oral cavity cancer: results of a large population-based case-control study, the ICARE study. *Cancer epidemiology*. 2013; 37: 284-9.
7. Fu JY, Gao J, Zhang ZY, et al. Tea consumption and the risk of oral cancer incidence: a case-control study from China. *Oral oncology*. 2013; 49: 918-22.
8. Chen F, He BC, Yan LJ, et al. Tea consumption and its interactions with tobacco smoking and alcohol drinking on oral cancer in southeast China. *Eur J Clin Nutr*. 2017; 71: 481-5.
9. He B, Chen F, Yan L, et al. Independent and joint exposure to passive smoking and cooking oil fumes on oral cancer in Chinese women: a hospital-based case-control study. *Acta oto-laryngologica*. 2016; 136: 1074-8.
10. Montero PH, Yu C, Palmer FL, et al. Nomograms for preoperative prediction of prognosis in patients with oral cavity squamous cell carcinoma. *Cancer*. 2014; 120: 214-21.
11. Chen F, Yan L, Lin L, et al. Dietary score and the risk of oral cancer: a case-control study in southeast China. *Oncotarget*. 2017; 8: 34610-6.
12. Kattan MW, Eastham JA, Stapleton AM, et al. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *Journal of the National Cancer Institute*. 1998; 90: 766-71.
13. Weiser MR, Landmann RG, Kattan MW, et al. Individualized prediction of colon cancer recurrence using a nomogram. *Journal of clinical oncology*. 2008; 26: 380-5.
14. Van Zee KJ, Manassseh DM, Bevilacqua JL, et al. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. *Annals of surgical oncology*. 2003; 10: 1140-51.
15. Bobdey S, Balasubramanian G, Mishra P. Nomogram prediction for survival of patients with oral cavity squamous cell carcinoma. *Head & neck*. 2016; 38: 1826-31.
16. Liu Y, Li Y, Fu Y, et al. Quantitative prediction of oral cancer risk in patients with oral leukoplakia. *Oncotarget*. 2017; 8: 46057-64.
17. Krishna Rao S, Mejia GC, Logan RM, et al. A screening model for oral cancer using risk scores: development and validation. *Community dentistry and oral epidemiology*. 2016; 44: 76-84.
18. Yan L, Chen F, He B, et al. A novel environmental exposure index and its interaction with familial susceptibility on oral cancer in non-smokers and non-drinkers: a case-control study. *European archives of oto-rhino-laryngology*. 2017; 274: 1945-50.
19. Daniel CR, Cross AJ, Graubard BI, et al. Prospective investigation of poultry and fish intake in relation to cancer risk. *Cancer prevention research*. 2011; 4: 1903-11.
20. MacLean CH, Newberry SJ, Mojica WA, et al. Effects of omega-3 fatty acids on cancer risk: a systematic review. *Jama*. 2006; 295: 403-15.
21. Ko KP, Park SK, Yang JJ, et al. Intake of soy products and other foods and gastric cancer risk: a prospective study. *Journal of epidemiology*. 2013; 23: 337-43.
22. Shin A, Lee J, Lee J, et al. Isoflavone and Soyfood Intake and Colorectal Cancer Risk: A Case-Control Study in Korea. *PloS one*. 2015; 10: e0143228.
23. Lee SA, Shu XO, Li H, et al. Adolescent and adult soy food intake and breast cancer risk: results from the Shanghai Women's Health Study. *The American journal of clinical nutrition*. 2009; 89: 1920-6.
24. Heck JE, Berthiller J, Vaccarella S, et al. Sexual behaviours and the risk of head and neck cancers: a pooled analysis in the International Head and Neck Cancer Epidemiology (INHANCE) consortium. *International journal of epidemiology*. 2010; 39: 166-81.
25. Anaya-Saavedra G, Ramirez-Amador V, Irigoyen-Camacho ME, et al. High association of human papillomavirus infection with oral cancer: a case-control study. *Archives of medical research*. 2008; 39: 189-97.
26. Chen F, Yan L, Liu F, et al. Oral human papillomavirus infection, sexual behaviors and risk of oral squamous cell carcinoma in southeast of China: A case-control study. *Journal of clinical virology*. 2016; 85: 7-12.