Risk factors for esophageal squamous cell carcinoma and its histological precursor lesions in China: a multicenter cross-sectional study

Chen Niu, Yong Liu, Jialin Wang, Yuqin Liu, Shaokai Zhang, Yongzhen Zhang, Liwei Zhang, Deli Zhao, Fugang Liu, Lina Chao, Xinzhen Wang, Chunli Zhang, Guohui Song, Zhiyi Zhang, Youpeng Li, Zheng Yan, Yongxiu Wen, Yinyin Ge, Zhaoping Zang, Wei Feng, Haiping Zhang, Lixin Tao, Rena Nakyeyune, Yi Shen, Yi Shao, Xiuhua Guo, Toni Miles, Aiming Yang, Fen Liu, and Guiqi Wang

Abstract

Background: Despite research efforts, the causative factors that contribute to esophageal squamous cell carcinoma (ESCC) in high-risk areas have not yet been understood. In this study, we, therefore, aimed to describe the risk factors associated with ESCC and its precursor lesions.

Methods: We performed an endoscopic examination of 44,857 individuals aged 40–69 years from five high incidence regions of China in 2017–2018. Participants were classified as 4 groups of normal control, esophagitis, low-grade intraepithelial neoplasia (LGIN) and high-grade intraepithelial neoplasia/esophageal squamous cell carcinoma (HGIN/ESCC) using an unconditional logistic regression determine risk factors.

Results: We identified 4890 esophagitis, 1874 LGIN and 437 HGIN/ESCC cases. Crude odds ratios (ORs) and adjusted odds ratios were calculated using unconditional logistic regression. Drinking well and surface water, salty diet, and positive family history of cancer were the common risk factors for esophagitis, LGIN and HGIN/ESCC. History of chronic hepatitis/cirrhosis was the greatest risk factor of esophagitis (adjusted OR 2.96, 95%CI 2.52–3.47) and HGIN/ESCC (adjusted OR 1.91, 95%CI 1.03–3.22). Pesticide exposure (adjusted OR 1.20, 95%CI 1.05–1.37) was essential risk factor of LGIN.

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Conclusions: Among individuals aged 40–69 years in high incidence regions of upper gastrointestinal cancer, the results provided important epidemiological evidence for the prevention of different precancerous lesions of ESCC.

Keywords: Esophageal squamous cell carcinoma, Precancerous lesions, Cross-sectional study, Risk factor, Prevention

Background
Esophageal cancer ranks seventh in terms of incidence and sixth in mortality overall according to the report of the 2018 Global Cancer Statistics [1]. In China, esophageal cancer is the sixth most frequent malignant tumor and the fourth leading cause of cancer-related death [2]. Esophageal adenocarcinoma (EADC) and esophageal squamous cell carcinoma (ESCC) are the two most common histologic subtypes of esophageal cancer [3]. In high-income countries, majority of the esophageal cancer patients have EADCs. In China, more than 90% of the esophageal cancer cases are ESCCs [4]. The most studied region of China is located in the North Central Taihang Mountain. In this area, ESCC is the leading cause of death [5]. Based on a recent report from this region, esophageal cancer prognosis is poor and the 5-year survival rates are about 30% [6].

Esophagitis is usually a benign disease. However, it is associated with an elevated risk for subsequent ESCC. Based on the WHO tumor classification, endoscopic biopsy of precancerous lesions yield low-grade intraepithelial neoplasia (LGIN) or high-grade intraepithelial neoplasia (HGIN) [7]. A report using a sample of containing both lesion types reported a 74% risk for esophageal cancer [8]. Identification of precursor lesions is critical for clinical care because these precursor lesions are curable. Early identification can reduce or delay ESCC mortality [9]. Overall, epidemiologic studies of ESCC show that it is population-dependent. The multifactorial basis for ESCC includes environmental exposures, lifestyle factors, and genetic traits [10]. In Western countries, alcohol consumption and smoking are the major risk factors for ESCC leading to a sevenfold increased risk of disease for smokers [11]. In high-risk areas of China, tobacco or alcohol does not appear to play the same role in population level rates of ESCC [12, 13]. Epidemiologic studies in these areas have not identified alternatives to smoking and alcohol consumption to explain these high rates of ESCC [14]. By increasing the number of cases of different esophageal lesions, this report presents evidence to improve our understanding of population-level factors in high rate areas of China.

To fill the gap in our understanding of risk factors at different stages of esophageal lesions - esophagitis, LGIN and HGIN - this report presents cross-sectional data from five high-incidence regions of ESCC in China.

Methods
Study population
This was a multicenter cross-sectional study, which relies on the high incidence regions of esophageal, gastric and colorectal cancer established by the cancer early diagnosis and early treatment project in China [15]. In 2017, we launched a new screening study of upper gastrointestinal cancer in five high-risk regions of upper gastrointestinal cancer in China, including Hebei, Henan, Shandong, Shanxi, and Gansu Provinces. The main purpose of this project is to identify the high-risk population of malignant upper gastrointestinal cancer and to establish a cancer risk prediction model to provide support for the prevention of upper gastrointestinal cancer.

The inclusion criteria were as follows: (1) local permanent residents in selected regions, (2) no history of endoscopic examination during the last 3 years, (3) no history of cancer, mental disorder, or any contraindication for endoscopy, (4) signed informed consent, and (5) agreement to complete all survey and examination, including endoscopy. The participant selection process is shown in Fig. 1. We recruited participants from April 2017 to December 2018. The final analysis included 44,857 residents aged 40–69 years. Among these participants, there were 116 invasive ESCC cases, 321 HGIN cases, 1874 LGIN cases, 4890 esophagitis cases, and 37,656 normal esophagus controls.

The study was approved by the Capital Medical University, Chinese Academy of Medical Sciences and Peking Union Medical College. All the experiment protocol for involving humans was in accordance to guidelines of Declaration of Helsinki. The STROBE statement can be found in Additional File 1.

Survey and measurement
Comprehensive questionnaire information was collected by face-to-face interviews and entered directly into a laptop based data entry system by trained investigators. The data entry process was facilitated with software designed to decrease missing items and reduce logic inaccuracy. A full questionnaire commonly took 35–45 min to accomplish. Items in the survey included demographic characteristics (e.g. gender, age), socio-economic status (e.g. occupation, education and income), lifestyle habits (e.g. smoking, alcohol and tea drinking), eating habits (e.g. fresh fruit, meat and milk), history of diseases (e.g. history of chronic hepatitis/cirrhosis), as well as...
medical history (e.g. acid suppressant medications and antibiotics). In this study, occupation and education level were also classified two categories: farmer/non-farmer and formal/no formal. In addition, frequency of consumption of specific dietary components was collected using an item with a five-level response category. The responses included: eating every day, 4–6 days per week, 1–3 days per week, 1–3 days per month and never eating. Body mass index (BMI) was calculated by the formula: weight (kg)/height (m^2) and then classified as underweight if BMI < 18.5 kg/m^2, normal if between 18.5 kg/m^2 and 23.9 kg/m^2, overweight if between 24.0 kg/m^2 and 27.9 kg/m^2, and obese if BMI ≥ 28.0 kg/m^2 according to the Chinese population recommendation [16]. The detailed definitions of the variables were shown in Additional File 2.

The endoscopic examinations were conducted by physicians at local hospitals. Procedures were based on clinical guidelines for cancer screening and early diagnosis and treatment in China [17]. The whole esophagus was visually examined. Lugol’ iodine staining in esophagus was applied to the diagnosis of suspicious lesions. The subjects with suspicious lesions (unstained and inflammatory lesions) were needed biopsy. The number of biopsies depended on the size of the lesion (>1 cm, ≥ 2 biopsy; >2 cm, ≥ 3 biopsy; >3 cm, ≥ 4 biopsy). In addition, if multiple scattered lesions were found, biopsy of all suspicious lesions should be taken as much as possible. Subjects without suspicious lesions did not receive a biopsy. The biopsy slides were read by two pathologists independently. If there were any inconsistencies, a third pathologist would give advice through discussion. Doctors reported the worst biopsy diagnosis to participants with multiple lesions. The histological criteria were as previously described [18–20].

To ascertain severity, esophageal mucosa was ranked using 5 categories: normal esophageal mucosa, minor mucosa changes, esophagitis, esophageal squamous simple hyperplasia (ESSH) or esophageal squamous dysplasia (ESD) [21]. ESD was further categorized into 3 levels including mild, moderate, and severe. According to WHO tumor histological classification [7], mild and moderate ESD combined fall under LGIN. Severe ESD and squamous cell carcinoma in situ (CIS) are considered as the HGIN. In this study, we grouped the participants into 4 groups: normal control, esophagitis, LGIN and HGIN/ESCC (lesions including severe ESD, CIS, and higher-grade lesions).

**Statistical analyses**

Basic, descriptive statistics show categorical variables as frequencies and percentages, while continuous variables are shown as mean and standard deviations. The appropriate tests for significance were applied, $\chi^2$ or ANOVA.
test, respectively. In this report, we use the unconditional univariable logistic model to calculate crude odds ratios (ORs) and 95% confidence intervals (CIs) [22]. The variables with a P value ≤ 0.2 in unconditional univariable logistic analysis were selected for unconditional multivariable logistic analysis. We considered three models to calculate the adjusted OR, including age-adjusted only, gender-adjusted only and adjusted age, gender, education, BMI and income at the same time. These adjusted confounders were selected based on the previous studies of esophageal cancer. Besides, models of adjustment for age or gender only were conducted as they may influence the socioeconomic status. In this study, dependent variable was the diagnosis outcome of disease. The independent variables included environment, lifestyle, diet habits, and medical history. We used SPSS software (version 24.0) to perform $\chi^2$ or ANOVA test, and glm function of R software (version 3.6.1) to perform logistic regression analysis. All tests of significance were two-sided with a P value of 0.05 was considered as statistically significant, expect univariable logistic regression analysis.

### Results

#### Participant characteristics

A total of 44,857 subjects aged 40 to 69 were incorporated in the analysis, including 37,656 (83.95%) subjects with a normal esophagus, 4890 (10.90%) people with esophagitis, 1874 (4.18%) cases of LGIN and 437 (0.97%) cases of HGIN/ESCC. Because the complete statistical description table was too long (Additional File 3), we only selected a few main variables to show in Table 1. The proportion of male in normal group, esophagitis, LGIN and HGIN/ESCC was 41.4, 46.8, 48.7 and 56.5%, respectively. However, the normal group (54.41 ± 7.26) was composed of a slightly younger population than the esophagitis (57.29 ± 6.85), LGIN (59.22 ± 6.25) and HGIN/ESCC (60.85 ± 5.80) group.

#### Risk factors for different esophageal lesions

We applied two steps to identify risk factors. Firstly, the unconditional univariable logistic regression analysis was used to test each variable independently and the estimated ORs with 95%CI are listed in Additional File 4. Secondly, the statistically significant variables in unconditional univariable logistic analysis were selected for unconditional multivariable analysis and the results were shown in Additional File 5. Tables 2, 3, 4 displayed the significant factors of 3 lesion groups from unconditional univariable logistic analysis and multivariable logistic analysis adjusted for age, gender, education, BMI and income.

After final unconditional multivariable analysis, we identified a few important risk factors of 3 esophageal lesions. For the HGIN/ESCC patients, some unhealthy lifestyle and dietary habits (Table 3) were risk factors, including alcohol (adjusted OR 1.60, 95%CI 1.23–2.09), smoking (adjusted OR 1.46, 95%CI 1.14–1.89), drinking

---

**Table 1** Characteristics of participants and endoscopy results of different esophageal lesions in Chinese high-risk population in 2017–2018

| Factors                  | HGIN/ESCC n = 437 | LGIN n = 1874 | Esophagitis n = 4890 | Norma esophagus n = 37,656 | Total cohort n = 44,857 | P value |
|--------------------------|-------------------|---------------|----------------------|-----------------------------|-------------------------|---------|
| Gender, n(%)             |                   |               |                      |                             |                         | <0.001a |
| Male                     | 247 (56.5)        | 913 (48.7)    | 2290 (46.8)          | 15,590 (41.4)               | 19,040 (42.4)           |         |
| Female                   | 190 (43.5)        | 961 (51.3)    | 2600 (53.2)          | 22,066 (58.6)               | 25,817 (57.6)           |         |
| Age, mean (SD)           | 60.85 (5.80)      | 59.22 (6.25)  | 57.29 (6.85)         | 54.41 (7.26)                | 54.99 (7.30)            | <0.001b |
| Education, n(%)          |                   |               |                      |                             |                         | <0.001a |
| No formal               | 81 (18.5)         | 368 (19.6)    | 744 (15.2)           | 5395 (14.3)                 | 6588 (14.7)             |         |
| Formal                  | 356 (81.5)        | 1506 (80.4)   | 4146 (84.8)          | 32,261 (85.7)               | 38,269 (85.3)           |         |
| Occupation, n(%)         |                   |               |                      |                             |                         | <0.001a |
| Non-farmer              | 86 (19.7)         | 382 (20.4)    | 1028 (21.0)          | 9347 (24.8)                 | 10,843 (24.2)           |         |
| Farmer                  | 351 (80.3)        | 1402 (79.6)   | 3862 (79.0)          | 28,309 (75.2)               | 34,014 (75.8)           |         |
| BMI (kg/m²), n(%)        |                   |               |                      |                             |                         | <0.001a |
| < 18.5                  | 16 (3.7)          | 36 (1.9)      | 98 (2.0)             | 714 (1.9)                   | 864 (1.9)               |         |
| 18.5 to 23.9            | 229 (52.3)        | 914 (48.8)    | 2336 (47.8)          | 17,309 (46.0)               | 20,788 (46.3)           |         |
| 24.0 to 27.9            | 159 (36.4)        | 742 (39.6)    | 1977 (40.4)          | 15,074 (40.0)               | 17,952 (40.1)           |         |
| ≥ 28.0                  | 33 (7.6)          | 182 (9.7)     | 479 (9.8)            | 4559 (12.1)                 | 5253 (11.7)             |         |

Source: WHO tumor histological classification 2000. * = $\chi^2$ test; ** = ANOVA test. Population recruited from five high-rate regions for upper gastrointestinal cancer in China, including Hebei, Henan, Shandong, Shanxi, and Gansu Provinces

BMI body mass index (kg/m²), LGIN low-grade intraepithelial neoplasia, HGIN high-grade intraepithelial neoplasia, ESCC esophageal squamous cell carcinoma
Table 2 Demographic characteristics associated with different esophageal lesions

| Demographic characteristics | HGIN/ESCC | P value | Adjusted OR a (95% CI) | LGIN | P value | Adjusted OR a (95% CI) | Esophagitis | P value | Adjusted OR a (95% CI) |
|-----------------------------|----------|---------|------------------------|------|---------|------------------------|-------------|---------|------------------------|
| Gender (Female)             | 0.54 (0.45–0.66) | <0.001 | 1.00 (1.03–1.06) | 0.74 (0.68–0.82) | <0.001 | 0.80 (0.76–0.85) | <0.001 |
| Age, years                  | 40–49    | 1.00 (reference) | <0.001 | 1.00 (reference) | 1.00 (reference) | 1.00 (reference) | <0.001 |
| 50–59                       | 4.96 (3.12–8.40) | <0.001 | 3.17 (2.65–3.81) | <0.001 | 1.80 (1.65–1.96) | <0.001 | <0.001 |<0.001 |
| 60–69                       | 15.73 (10.06–26.29) | <0.001 | 7.18 (6.04–8.60) | <0.001 | 2.84 (2.60–3.10) | <0.001 | <0.001 |<0.001 |
| Education (Formal)          | 0.73 (0.58–0.94) | 0.013 | 0.68 (0.61–0.77) | <0.001 | 0.93 (0.86–1.01) | <0.001 | <0.001 |<0.001 |
| Annual income per family (RMB) | <0.001 |<0.001 |1.00 (reference) | <0.001 | 1.00 (reference) | <0.001 | <0.001 |<0.001 |
| <10,000                     | 0.65 (0.49–0.89) | 0.005 | 1.00 (0.86–1.18) | <0.001 | 0.86 (0.78–0.95) | <0.001 | <0.001 |<0.001 |
| 10,000–14,999               | 0.80 (0.61–1.07) | 0.124 | 0.725 (0.88–1.20) | <0.001 | 1.02 (0.93–1.13) | <0.001 | <0.001 |<0.001 |
| 50,000–74,999               | 0.60 (0.43–0.85) | 0.004 | 0.001 |<0.001 |<0.001 |<0.001 |<0.001 |<0.001 |
| ≥70,000                     | 0.42 (0.26–0.65) | <0.001 | 0.58 (0.46–0.73) | <0.001 | 1.26 (0.99–1.61) | <0.001 | <0.001 |<0.001 |
| BMI                         | 1.00 (reference) | <0.001 | 1.00 (reference) | <0.001 | 1.00 (reference) | <0.001 | <0.001 |<0.001 |
| 18.5 to 23.9                | 0.69 (0.97–2.73) | 0.044 | 0.95 (0.67–1.32) | <0.001 | 1.02 (0.82–1.26) | <0.001 | <0.001 |<0.001 |
| 24.0 to 27.9                | 0.80 (0.65–0.98) | 0.029 | 0.90 (0.84–1.03) | <0.001 | 0.97 (0.91–1.04) | <0.001 | <0.001 |<0.001 |
| ≥28.0                       | 0.55 (0.37–0.78) | 0.001 | 0.76 (0.64–0.89) | 0.001 | 0.78 (0.70–0.86) | <0.001 | <0.001 |<0.001 |
| Family history of cancer    | 1.50 (1.23–1.82) | 0.001 | 1.64 (1.34–2.00) | <0.001 | 1.24 (1.12–1.37) | <0.001 | 1.33 (1.20–1.47) | <0.001 |
| History of chronic hepatitis and cirrhosis | 1.96 (1.07–3.28) | 0.018 | 1.91 (1.03–3.22) | 0.025 | 1.36 (0.96–1.86) | 0.067 | 1.36 (0.96–1.87) | 0.070 |

Source: WHO tumor histological classification 2000
BMI body mass index (kg/m²), LGIN low-grade intraepithelial neoplasia, HGIN high-grade intraepithelial neoplasia, ESCC esophageal squamous cell carcinoma
*adjusted for age, gender, education, BMI and income

well and surface water (adjusted OR 1.55, 95% CI 1.28–1.89), salty diet (adjusted OR 1.36, 95% CI 1.07–1.77), drinking tea (adjusted OR 1.32, 95% CI 1.08–1.61). Although the P value of drinking hot/burning hot tea was no statistically significant, the adjusted OR value was 1.34. In addition, pesticide exposure (adjusted OR 1.44, 95% CI 1.11–1.84), loose teeth (adjusted OR 1.48, 95% CI 0.99–2.12), history of chronic hepatitis/cirrhosis (adjusted OR 1.91, 95% CI 1.03–3.22), and family history of cancer (adjusted OR 1.64, 95% CI 1.34–2.00), also positively related to the risk of HGIN/ESCC (Tables 2, 3, 4). Based on the results of the adjusted analysis, we also found that some risk factors associated with LGIN were the same as HGIN/ESCC, consisting of drinking alcohol (adjusted OR 1.30, 95% CI 1.12–1.51), drinking well and surface water (adjusted OR 1.19, 95% CI 1.08–1.32), salty diet (adjusted OR 1.57, 95% CI 1.38–1.81), drinking tea (adjusted OR 1.34, 95% CI 1.26–1.54), family history of
| Environment, lifestyle and habits | HGIN/ESCC | LGIN | Esophagitis |
|----------------------------------|-----------|------|-------------|
|                                  | OR (95% CI) | P value | Adjusted OR * (95% CI) | OR (95% CI) | P value | Adjusted OR * (95% CI) | OR (95% CI) | P value | Adjusted OR * (95% CI) |
| **Smoking**                      |           |       |             |           |       |             |           |       |             |
| Not smoke                        | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – |
| Former/current smoke             | 1.97 (1.62–2.41) | < 0.001 | 1.46 (1.14–1.89) | 0.03 | 1.28 (1.15–1.43) | < 0.001 | 1.04 (0.91–1.19) | 0.574 | 1.25 (1.16–1.33) | < 0.001 |
| **Occupation (Farmer versus Non-farmer)** | 1.35 (1.07–1.72) | 0.014 | 1.22 (0.97–1.57) | 0.31 | 1.29 (1.15–1.45) | < 0.001 | 1.20 (1.07–1.35) | 0.002 | 1.24 (1.15–1.33) | < 0.001 |
| Pesticide exposure               | 1.46 (1.13–1.87) | 0.003 | 1.44 (1.11–1.84) | 0.05 | 1.23 (1.08–1.40) | < 0.001 | 1.20 (1.05–1.37) | 0.007 | 1.04 (0.95–1.13) | – |
| **Refrigerator use, length of time** |           |       |             |           |       |             |           |       |             |
| No refrigerator                  | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – |
| 1-10 years                       | 0.57 (0.43–0.78) | 0.001 | 0.76 (0.57–1.04) | 0.074 | 0.74 (0.63–0.87) | < 0.001 | 0.88 (0.75–1.04) | 0.135 | 0.66 (0.60–0.74) | < 0.001 |
| 11-20 years                      | 0.52 (0.36–0.76) | 0.011 | 0.71 (0.49–1.05) | 0.087 | 0.64 (0.52–0.78) | < 0.001 | 0.78 (0.64–0.95) | 0.015 | 0.61 (0.54–0.69) | < 0.001 |
| > 20 years                       | 0.44 (0.11–1.21) | 0.170 | 0.55 (0.13–1.52) | 0.319 | 0.60 (0.33–1.01) | < 0.001 | 0.69 (0.37–1.16) | 0.189 | 0.48 (0.32–0.69) | 0.001 |
| **Alcohol**                      |           |       |             |           |       |             |           |       |             |
| Not drink                        | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – |
| Former/current drink             | 2.15 (1.68–2.75) | < 0.001 | 1.60 (1.23–2.09) | 0.001 | 1.54 (1.34–1.76) | < 0.001 | 1.30 (1.12–1.51) | < 0.001 | 1.22 (1.10–1.34) | < 0.001 |
| **Tea drinking frequency**       |           |       |             |           |       |             |           |       |             |
| Not drink                        | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – |
| Former/current drink             | 1.35 (1.11–1.65) | 0.003 | 1.32 (1.08–1.61) | 0.007 | 1.43 (1.30–1.57) | < 0.001 | 1.34 (1.26–1.54) | < 0.001 | 0.78 (0.73–0.83) | < 0.001 |
| Water source (Well water and surface water) | 1.64 (1.35–1.98) | 0.001 | 1.55 (1.28–1.89) | 0.001 | 1.25 (1.13–1.37) | < 0.001 | 1.19 (1.08–1.32) | < 0.001 | 1.59 (1.50–1.69) | 0.001 |
| Tea temperature *                |           |       |             |           |       |             |           |       |             |
| Warm tea                         | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – |
| Hot/burning hot tea              | 1.34 (0.91–1.99) | 0.144 | 1.34 (0.90–1.99) | 0.151 | 1.11 (0.92–1.33) | 0.277 | 1.10 (0.92–1.33) | 0.300 | 1.10 (0.95–1.27) | 0.202 |
| Drink improved water             | 0.74 (0.49–1.08) | 0.137 | 0.79 (0.52–1.15) | 0.251 | 0.56 (0.44–0.69) | < 0.001 | 0.58 (0.46–0.72) | < 0.001 | 0.72 (0.64–0.82) | 0.001 |
| **Diet taste**                   |           |       |             |           |       |             |           |       |             |
| Light diet                       | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – |
| Salty diet                       | 1.40 (1.09–1.81) | 0.009 | 1.36 (1.07–1.77) | 0.016 | 1.60 (1.41–1.83) | < 0.001 | 1.57 (1.38–1.80) | < 0.001 | 1.59 (1.47–1.74) | 0.001 |
| **Livestock meat**               | 0.90 (0.71–1.18) | 0.440 | – | – | 1.00 (0.88–1.14) | – | 0.83 (0.76–0.89) | < 0.001 | 0.88 (0.81–0.95) | 0.001 |
| Poultry meat                     | 0.77 (0.63–0.95) | 0.014 | 0.89 (0.72–1.10) | 0.290 | 0.66 (0.59–0.73) | < 0.001 | 0.73 (0.66–0.81) | < 0.001 | 0.59 (0.56–0.64) | 0.001 |
| Seafood                          | 0.99 (0.77–1.27) | 0.967 | – | – | 1.08 (0.95–1.21) | 0.220 | – | – | 0.86 (0.80–0.94) | 0.001 |
| Fruits                           | 0.78 (0.73–0.83) | 0.084 | 0.085 | 0.79 (0.71–0.87) | 0.015 | 0.70 (0.58–0.85) | < 0.001 | 0.83 (0.78–0.88) | 0.001 |

*Adjusted OR: Adjusted odds ratio. CI: confidence interval. *P < 0.05 is considered statistically significant.
cancer (adjusted OR 1.33, 95% CI 1.20–1.47), pesticide exposure (adjusted OR 1.20, 95% CI 1.05–1.37) and loose teeth (adjusted OR 1.23, 95% CI 1.00–1.51). In addition, occupation as a farmer (adjusted OR 1.20, 95% CI 1.07–1.35) was also positively related to the risk of LGIN (Table 3). Drinking hot tea still did not find statistical significance (adjusted OR 1.10, 95% CI 0.92–1.33). On the contrary, eating poultry meat (adjusted OR 0.73, 95% CI 0.66–0.81), fruits (adjusted OR 0.89, 95% CI 0.80–0.98), vitamins (adjusted OR 0.49, 95% CI 0.27–0.82), milk (adjusted OR 0.82, 95% CI 0.71–0.94), acid suppressants (adjusted OR 0.51, 95% CI 0.39–0.66), antibiotics (adjusted OR 0.33, 95% CI 0.14–0.80) and drinking improved water (adjusted OR 0.58, 95% CI 0.46–0.72) were negatively related to the risk of LGIN.

Although esophagitis is a benign disease, there are more risk elevating and risk reduction factors than the diseases mentioned above. For the diet habits (Table 3), except that leftovers (adjusted OR 1.49, 95% CI 1.40–1.58) and salty diet (adjusted OR 1.57, 95% CI 1.45–1.71) were risk factors, the remaining factors had protective effect on esophagitis, such as poultry meat (adjusted OR 0.62, 95% CI 0.58–0.67), fruits (adjusted OR 0.83, 95% CI 0.78–0.88), vitamins (adjusted OR 0.68, 95% CI 0.50–0.90), livestock meat (adjusted OR 0.88, 95% CI 0.81–0.95), seafood (adjusted OR 0.89, 95% CI 0.82–0.97), spring onion/ginger/garlic (adjusted OR 0.69, 95% CI 0.65–0.74) and nuts (adjusted OR 0.77, 95% CI 0.71–0.84).

In addition, some unhealthy lifestyles were positively related to the risk of esophagitis (Table 3), including smoking (adjusted OR 1.09, 95% CI 1.00–1.19), drinking well and surface water (adjusted OR 1.61, 95% CI 1.51–1.71), and having snore (adjusted OR 1.09, 95% CI 1.02–1.15). On the contrary, other healthy lifestyles were negatively related to the risk of esophagitis (Table 3), including drinking improved water (adjusted OR 0.73, 95% CI 0.64–0.83), refrigerator use (adjusted OR 0.50–0.73, P < 0.001), physical exercise (adjusted OR 0.69, 95% CI 0.62–0.76) and doing housework (adjusted OR 0.57–0.63, P < 0.001). Except for foods and lifestyles, family history of cancer (adjusted OR 1.22, 95% CI 1.14–
Table 4 Medical history associated with different esophageal lesions

| Medical history                        | HGIN/ESCC  | LGIN  | Esophagitis |
|----------------------------------------|------------|-------|-------------|
|                                        | OR (95% CI)| P value| OR (95% CI)| P value| OR (95% CI)| P value| OR (95% CI)| P value|
| Take an acid suppressant               |            |       |            |       |            |       |            |       |
|                                          | 0.65 (0.39–1.02) | 0.077 | 0.68 (0.41–1.07) | 0.119 | 0.49 (0.37–0.63) | < 0.001 | 0.51 (0.39–0.66) | < 0.001 | 0.68 (0.59–0.79) | < 0.001 | 0.70 (0.60–0.80) | < 0.001 |
| Take antibiotics                        |            |       |            |       |            |       |            |       |
| Not take                               | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – |
| Not every week                          | 0.27 (0.07–1.08) | 0.063 | 0.34 (0.08–1.36) | 0.127 | 0.76 (0.12–1.72) | 0.180 | 0.90 (0.59–1.36) | 0.617 | 0.54 (0.40–0.73) | < 0.001 | 0.61 (0.45–0.83) | < 0.001 |
| Not every day                           | 0.68 (0.17–2.74) | 0.588 | 0.81 (0.20–3.28) | 0.764 | 0.80 (0.43–1.51) | 0.492 | 0.93 (0.49–1.77) | 0.826 | 0.70 (0.461.08) | 0.107 | 0.75 (0.49–1.16) | 0.197 |
| Every day                               | 0.25 (0.04–1.80) | 0.169 | 0.28 (0.04–1.99) | 0.202 | 0.30 (0.12–0.72) | 0.007 | 0.33 (0.14–0.80) | 0.014 | 0.39 (0.24–0.63) | < 0.001 | 0.40 (0.24–0.65) | < 0.001 |

Number of teeth lost

|                                        |            |       |            |       |            |       |            |       |
| None                                   | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.00 (reference) | – | 1.54 (1.14–1.22) | < 0.001 | 1.54 (1.14–1.45) | < 0.001 |
| 1–3                                    | 1.07 (0.84–1.36) | 0.570 | 0.86 (0.68–1.09) | 0.210 | 1.38 (1.24–1.54) | < 0.001 | 1.17 (1.04–1.31) | 0.007 | 1.24 (1.16–1.33) | < 0.001 | 1.14 (1.06–1.23) | < 0.001 |
| 4–6                                    | 1.68 (1.26–2.22) | < 0.001 | 1.11 (0.83–1.48) | 0.484 | 1.53 (1.31–1.77) | < 0.001 | 1.12 (0.96–1.30) | 0.137 | 1.35 (1.22–1.48) | < 0.001 | 1.15 (1.04–1.27) | < 0.005 |
| 7–11                                   | 1.54 (1.01–2.26) | 0.035 | 0.83 (0.55–1.25) | 0.382 | 1.88 (1.55–2.26) | < 0.001 | 1.19 (0.98–1.44) | 0.077 | 1.61 (1.41–1.82) | < 0.001 | 1.27 (1.12–1.45) | < 0.001 |
| 12–31                                  | 2.62 (1.70–3.87) | 0.001 | 1.21 (0.79–1.84) | 0.385 | 2.18 (1.73–2.72) | < 0.001 | 1.26 (0.99–1.58) | 0.054 | 2.05 (1.76–2.37) | < 0.001 | 1.54 (1.32–1.79) | < 0.001 |
| Complete denture                       | 2.72 (1.72–4.10) | < 0.001 | 1.20 (0.77–1.87) | 0.420 | 2.55 (2.02–3.19) | < 0.001 | 1.42 (1.11–1.78) | 0.004 | 2.66 (2.30–3.07) | < 0.001 | 1.94 (1.66–2.25) | < 0.001 |
| Loose teeth                            | 1.89 (1.27–2.70) | 0.001 | 1.48 (0.99–2.12) | 0.042 | 1.49 (1.21–1.82) | < 0.001 | 1.23 (1.00–1.51) | 0.049 | 1.23 (1.06–1.42) | 0.005 | 1.10 (0.95–1.27) | 0.202 |

Source: WHO tumor histological classification 2000

BMI body mass index (kg/m²), LGIN low-grade intraepithelial neoplasia, HGIN high-grade intraepithelial neoplasia, ESCC esophageal squamous cell carcinoma
*aadjusted for age, gender, education, BMI and income

1.30), history of chronic hepatitis/cirrhosis (adjusted OR 2.96, 95%CI 2.52–3.47), occupation as a farmer (adjusted OR 1.21, 95%CI 1.12–1.30) and missing teeth (adjusted OR 1.14–1.94, P < 0.01) were positively related to the risk of esophagitis (Tables 2, 3, 4). Taking acid suppressants (adjusted OR 0.70, 95%CI 0.60–0.80) and taking antibiotics every day (adjusted OR 0.40, 95%CI 0.24–0.65) were negatively related to the risk of esophagitis (Table 4).

Discussion

The high rate of early esophageal lesions progressing to esophageal cancer has resulted in a concerted attempt to better understand the risk factors associated with this difference. In this study, we conducted a cross-sectional study in five areas with high incidence of ESCC in China, aiming to recognize the risk factors related to different esophageal lesions. At last, we identified three important risk factors, which persisted across all stages. Otherwise, we also find that the key risk factors in different stages were various.

The three common risk factors are drinking well and surface water, salty diet, and family history of cancer. At present, the main sources of drinking water for some local residents in rural area of high incidence region were still well and surface water. Since the 1980s, the local government has taken significant measures to improve the quality of drinking water, and the incidence and mortality of esophageal cancer displayed a decreasing trend compared with regions where no improvement [23]. Our result showed that drinking well and surface water was important risk factor from early stage to late stage of lesions. We advised that the water improvement project still needed to implement in rural regions with poor drinking water quality.

Previous studies found local people of rural areas were inclined to eat salty foods, such as salted meat, salted duck eggs and pickled vegetables. Both high frequency and long-term intakes were associated with elevated risk
of ESCC. This observation is consistent with other studies [16, 24, 25]. However, the underlying mechanism was still unclear. We speculate that the mechanism, possible similar to gastric cancer [26, 27], is that high salt concentrations disrupt the mucosal barrier and lead to inflammation, which could increase the possibility of bacterial infection. It is noticed that salty diet, an important risk factor, existed at all stages of early esophageal lesions in our study.

The association between ESCC and family history of cancer had been reported. Some studies showed that family history of cancer was strongly associated with ESCC risk and that genetic factors could contribute to the effect [28, 29]. Based on our results, we suggest that the above three common risk factors should be considered first in prevention of all stages.

History of chronic hepatitis/cirrhosis, pesticides exposure and alcohol were important risk factors for esophagitis, LGIN and HGIN/ESCC, respectively. China has a high burden of liver disease [30, 31], previous studies report that chronic liver diseases, in particular chronic hepatitis B virus infection, are associated with reflux esophagitis reflux [32, 33]. In this study, we identified that history of chronic hepatitis/cirrhosis was the tremendous risk factor both esophagitis and HGIN/ESCC. Currently, there is an effective vaccine to prevent hepatitis B virus infection. Increased vaccination for hepatitis B could also reduce the occurrence of ESCC in high incidence regions. Similar to our findings, previous studies had confirmed that alcohol is a risk factor for ESCC [4, 10]. The reason is that acetaldehyde, a class 1 carcinogen, is the first metabolite of ethanol metabolism [34, 35].

Pesticides had been reported as a risk factor for a variety of cancers [36], such as prostate cancer [37] and childhood leukemia [38], but the role of pesticides in esophageal cancer is still unclear. Our results showed that pesticide exposure was an important risk factor, which not only increased the risk of LGIN (adjusted OR 1.20, 95%CI 1.05–1.37) but also HGIN/ESCC (adjusted OR 1.44, 95% CI 1.11–1.84). In this study, most of participants are farmer, who had a greater chance exposed pesticide than non-farmer, which could increase the risk of early esophageal lesions. Therefore, the result showed that pesticide was a significant risk factor of LGIN and HGIN/ESCC. Further study is needed to evaluate the mechanism linking pesticide exposure and early esophageal lesions.

It is noticed that our study found that high frequency of drinking tea is a risk factor for LGIN and LGIN/ESCC, which is similar to previous study [39]. Currently, the association between tea consumption and esophageal cancer is inconsistent, which could be related to the type of tea [40]. In addition, our research found that although adjusted OR values were all greater than 1, there was no statistically significant correlation between high tea temperature and different esophageal lesions. This result could be caused by the small sample size of cases. Further prospective study is still needed. Interestingly, our results revealed that antibiotics were protective factor for esophagitis and LGIN. Previous study found that Porphyromonas gingivalis is positively correlated with the severity of esophageal squamous cell carcinoma and might contribute to the pathogenesis of esophageal squamous cell carcinoma [41]. We speculated that antibiotics intake might change the composition of the gut microbiome and reduce their adverse effects on esophagus, thus play a positive role in the prevention of ESCC. Besides, we also found that with the aggravation of the esophageal lesions, the number of risk factors reduced, which suggested that the effect of environmental factors on the late stage of the disease decreased. These results also mean that it is a more meaningful method for prevention of ESCC by controlling environmental factors in early lesion, such as balanced diet, proper exercise, and healthy lifestyle.

The present study had several limitations. Most of the variables analyzed in the study were based on self-reported data which may be susceptible to recall bias. In addition, there are some factors were not collected in the questionnaire, such as cooking methods, history of upper gastrointestinal tumors or esophageal cancer. The main strength of our study is that it was the first multicenter cross-sectional study to identify risk factors in different stage of early esophageal lesions. Furthermore, detailed questionnaire was collected in a standardized data entry system by trained investigator to assure the quality of the data. In addition, we used the gold standard method to exam esophageal lesions.

**Conclusions**

In conclusion, we performed a cross sectional study of 44,857 individuals in the age range of 40–69 years. These individuals live in regions with high-rates of ESCC regions in China. This report has identified drinking well and surface water, salty diet, family history of cancer and alcohol consumption as factors influencing disease rates across all stages. We also identified history of chronic hepatitis/cirrhosis and pesticide exposure as essential risk factors in esophagitis, LGIN and HGIN/ESCC, respectively. The results make a compelling case for environmental factors being the main driver of disease. The results also suggest directions for efforts to ESCC in an early stage of disease in high incidence regions. However, considering the complexity of the disease, although environmental factors play an important role in the early stages of ESCC, these may vary as the severity of the disease increases. Further cohort studies are needed to
support and expand our understanding of ESCC risk factors.

Abbreviations
BMI: Body mass index; CIs: Confidence intervals; CIS: Carcinoma in situ; EADC: Esophageal adenocarcinoma; ESCC: Esophageal squamous cell carcinoma; ESD: Esophageal squamous dysplasia; ESSH: Esophageal squamous simple hyperplasia; HGNI: High-grade intraepithelial neoplasia; LGNI: Low-grade intraepithelial neoplasia; ORs: Odds ratios

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12885-021-08764-x.

Acknowledgements
Not applicable

Authors’ contributions
Guqi Wang, Fen Liu, Aiming Yang and Xuhua Guo designed this study, Yong Liu, Jialin Wang, Yuqin Liu, Shaokai Zhang, Yongzhen Zhang, Liwei Zhong, Deli Zhao, Fugang Liu, Lina Chao, Xinzheng Wang, Chunli Zhang, Guohui Song, Zhiyi Zhang, Youpeng Li, Zheng Yan, Yongxiu Wen and Yinrui Ge contributed to the data collection and organized the database. Chen Niu, Zhaoping Zang, Wei Feng, Haiqing Zhang and Lixin Tao analyzed the data. Chen Niu draft the manuscript, Fen Liu, Renakaeyune, Yi Shen, Yi Shao and Toni Miles revised the manuscript. Guqi Wang, Fen Liu and Aiming Yang made the decision to submit for publication. The author(s) read and approved the final manuscript.

Funding
This work was supported by the National Key R&D Program of China (grant numbers 2016YFC1302804, 2016YFC1302800, 2016YFC0901400); National Natural Science Foundation of China (grant numbers 8174277, 8194493); CAMS Innovation Fund for Medical Sciences (grant numbers 2016I2M-1-001, 2017I2M-1-001, 2019I2M-2-004); and Sanming Project of Medicine in CAMS Innovation Fund for Medical Sciences (grant numbers 2016-I2M-1-001, 2019-I2M-2-004); and Sanming Project of Medicine in Shen-
cultural factors. Int J Hyg Environ Health. 2015;218(1):12–8. https://doi.org/10.1016/j.ijheh.2014.10.004.

15. Zou XN. Epidemic trend, screening, and early detection and treatment of cancer in Chinese population. Cancer Biol Med. 2017;14(1):50–9. https://doi.org/10.20892/j.issn.2095-3941.2016.0047.

16. Pan D, Su M, Zhang T, Mao C, Fu L, Yang L, et al. A distinct epidemiologic pattern of precancerous lesions of esophageal squamous cell carcinoma in a high-risk area of Huai’an, Jiangsu Province, China. Cancer Prev Res (Phila). 2011;4(7):449–62.

17. Dong Z. Guidelines for cancer screening and early diagnosis and treatment in China. Beijing: Peking University Medical Press; 2005.

18. Dawsey SM, Lewin KJ, Liu FS, Wang GQ, Shen Q. Esophageal morphology from Linxian, China. Squamous histologic findings in 754 patients. Cancer. 1994;73(8):2027–37. https://doi.org/10.1002/1097-0142(19940415)73:8<2027::AID-CNCR2820730803>3.0.CO;2-3.

19. Zeng H, Sun K, Cao M, Zheng R, Sun X, Liu S, et al. Initial results from a multi-center population-based cluster randomized trial of esophageal and gastric cancer screening in China. BMC Gastroenterol. 2020;20(1):398. https://doi.org/10.1186/s12876-020-01517-3.

20. Wei WQ, Hao OQ, Quan CT, Song GH, Wang M, Zhao DL, et al. Esophageal histological precursor lesions and subsequent 8.5-year cancer risk in a population-based prospective study in China. Am J Gastroenterol. 2020;115(7):1036–44. https://doi.org/10.14309/ajg.0000000000002640.

21. Liu X, Zhang M, Ying S, Zhang C, Lin R, Zheng J, et al. Genetic alterations in esophageal tissues from squamous dysplasia to carcinoma. Gastroenterology. 2017;153(1):166–77. https://doi.org/10.1053/j.gastro.2017.03.033.

22. Kuo CL, Duan Y, Grady J. Unconditional or conditional logistic regression model for age-matched case-control data? Front Public Health. 2018;6:57. https://doi.org/10.3389/fpubh.2018.00057.

23. Han JY, Xu ZX, Xing HP. A historical cohort study on the effect of drinking water improvement to the mortality of esophageal carcinoma among Linzhou residents. Zhonghua Liu Xing Bing Xue Za Zhi. 2007;28(5):515–6.

24. Jessri M, Rashidkhani B, Hajizadeh B, Jessri M, Kreiger N, Bajdik CD. Smoking, alcohol drinking, green tea consumption and the risk of esophageal cancer in Japanese men. J Epidemiol. 2006;16(5):185–92. https://doi.org/10.1097/1533-9824.0000000000000658.

25. Baba Y, Iwatsuki M, Yoshida N, Watanabe M, Baba H. Review of the gut microbiome and esophageal cancer: pathogenesis and potential clinical implications. Ann Gastroenterol Surg. 2017;1(2):99–104. https://doi.org/10.1002/ags3.12014.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.