Tobacco consumption has interaction effect with protein coding genes to increase esophageal squamous cell Carcinoma risk: A case-control study in Chinese high-incidence region

Renjia Zhao
Fudan University
Huangbo Yuan
Fudan University
Zhenqiu Liu
Fudan University
Xiaoxiao Li
Fudan University
Huiyao Chen
Fudan University
Xiaorong Yang
Shandong University
Weimin Ye
Fudan University Taizhou Institute of Health Sciences
Ming Lu
Shandong University
Li Jin
Fudan University Taizhou Institute of Health Sciences
Tiejun Zhang
Fudan University
Xingdong Chen
Fudan University
Chen Suo (suochen@fudan.edu.cn)
Fudan University

Research Article

Keywords: Esophageal squamous cell carcinoma; Single nucleotide polymorphism; Interaction; Smoke, Genetic Risk Score

Posted Date: December 3rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1133833/v1

Abstract

Background: Esophageal squamous cell carcinoma (ESCC) has reported that smoking was a major risk factor. Genetic predisposition can partially explain the pathogenesis of esophageal cancer in alcohol drinkers. It would also be interesting to investigate the genetic basis underlying the significant disparities in ESCC risk in populations exposed to the same level of smoking.

Methods: We recruited 1030 ESCC patients and 1783 healthy individuals in Taixing, China, and selected 101 ESCC-related SNPs for analysis. Logistic regression model was employed with an interaction term for smoking and individual SNPs. The level of smoking was categorized based on pack-years as never smokers, moderate smokers (≤30), and heavy smokers (>30). The relative excess risk of interaction (RERI) and the synergy index (S) were used to evaluate interaction on an additive scale. Genetic risk score was established to quantify gene risk.

Results: The SNP HECTD4 rs11066280, CASP8 rs3769823 and ADH1B rs1042026 had an interaction effect with smoking on ESCC risk. Specifically, for rs11066280 T/A, the adjusted OR for moderate and heavy smokers was 1.43 (95% CI = 1.01-2.02) and 1.80 (95% CI = 1.28–2.53), ADH1B rs1042026 showed strong effect in both smoker and alcohol drinkers. GRS indicated that these three SNPs had unsignificant effect in non-smokers and a 2.92-fold risk (95%CI = 1.69-5.11) in smokers.

Discussion and Conclusion: In this study, we provide new insights for disease prevention and control of ESCC based on smoking behavior and genetic predisposition.

Introduction

Esophageal cancer is the sixth most common cancer and the eighth most common cause of cancer-related death worldwide(1).There are two major histological types—esophageal adenocarcinoma (EA) and esophageal squamous cell carcinoma (ESCC)—which are different in terms of both epidemiology and pathogenesis(2).ESCC accounts for over 90% of esophageal cancer cases in the world and has a high incidence rate in East Asia and developing countries(3). In China, it is responsible for nearly 250,000 deaths annually(4) and has a poor prognosis, with a reported 5-year survival rate of 20%(1, 5).

Diet and lifestyle can greatly influence the risk of developing ESCC, and tobacco smoking is one of the major lifestyle factors that increase the risk of ESCC(6–10). Since 2007, Freedman et.al(11) reported that in a prospective cohort study, they found smoking could cause more than nine times of risk on ESCC (HR=9.27, 95%CI: 4.04~21.29), while in other latter researches, two to ten times of risk were calculated in heavy smokers than in former or non-smokers to increase ESCC(12, 13).

Besides behavior-related factors, it is essential to investigate the genetic basis underlying the significant disparities in ESCC risk. In fact, genetic predisposition has been shown to be an important mechanism in the pathogenesis of ESCC(14–17)and has been found to influence the survival of ESCC patients(18, 19). Researchers have identified several suspectable single nucleotide polymorphisms (SNPs), which are the most common type of genetic variation, that are associated with the risk of ESCC(20, 21). The major five type of genes were found to increase most of cancers, especially ESCC risk, including X-ray repair cross complementing group 1(XRCC1)(22, 23),cytochrome P450(CYP450)(24), cell cycle gene(25), methylenetetrahydrofolate reductase (MTHFR)(26, 27) and alcohol-related metabolism gene(28). However, the interaction effect between smoking these genes were still limited.

So, in this study, we performed a large population-based case-control study to investigate the potential interactions between SNPs and cigarettes consumption exposure and their association with ESCC risk. The findings provide insights for risk stratification of the population and disease prevention strategies.

Materials And Methods

Patients and Sample Selection

This study population included 1083 ESCC cases and 1786 controls. All cases were diagnosed between October 2010 and September 2013 in Taxing, which is one of the cities with the highest incidence of ESCC in East China. More than 90% of the esophageal cancer patients in this area are referred to the 4 largest hospitals (the People's Hospital of Taixing, the Second People's Hospital of Taixing, the Third People's Hospital of Taixing, and the Hospital of Traditional Chinese Medicine of Taixing), individuals diagnosed by the endoscopy units in these hospitals were invited to participate. This approach was designed to reduce nondifferential recall bias, given that these patients were unaware of their cancer diagnosis at the time of recruitment and data collection.(29) All personal and clinical data were surveyed by trained staff with the help of a specifically designed electronic questionnaire. We also matched the cases and controls by age and sex. The detailed study design, including quality control and inclusion and exclusion criteria, have been previously reported(10, 30-32). The patient selection process is shown in **Figure 1**.

SNP Genotyping, Screening and Quality Control

For this study, 101 SNPs from 59 genes were selected. All SNPs have been previously reported as ESCC susceptibility loci and identified by Genome-wide association study. The SNPs were genotyped using a three-round multiplex polymerase chain reaction procedure with next-generation sequencing method.

To ensure genotyping accuracy, we also implemented quality control procedures, such as by including negative controls. In addition, a randomly selected 8% of total samples were genotyped twice and the consistency was higher than 98%. The average sequencing depth was 1225x. All SNPs had a minor allele frequency of 0.1 or more in both the case patient and control samples, rendering adequate statistical

power. Among the 101 SNPs, 4 SNPs were monozygotic, 14 did not reach Hardy–Weinberg equilibrium, and 5 had a missing rate of >10%.

Definition of Tobacco consumption

We collected detailed data on tobacco consumption, including smoking status (never, ever, and current smoker), cigarettes consumption exposure in pack-years, and deep inhalation during smoking (Yes/No). Moreover, according to National Cancer Institutes' recommendation(33-36), cigarettes consumption exposure was redefined in the form of categorical variables by 30 pack-years, namely, never smokers (0 pack-years), moderate smokers (≤30 pack-years), and heavy smokers (>30 pack-years).

Definition of other variables

In this case-control study, participants were interviewed face-to-face with structured questionnaires, and information on basic characteristics was collected, including age, sex, smoking, alcohol intake, education status, wealth score, marital status, hot tea consumption. Alcohol intake was measured in three categories. (Never/Quitted/Drinker). Wealth score was a continues indictor calculated by a multiple correspondence analysis. It was consisted of the ownership of valuable home items for each participant, like television, cars, washing machines, vacuum cleaner etc. We defined the hot tea or not by calculating the time between placing tea leaves mixed with boiling water and tea drinking. Time less than 5 minutes was regarded as hot tea (Yes), then the warm tea (No).

Statistical analysis

Student *t*-test was used to analyze differences in continuous variables between the case and control groups; the chi-square test, to analyze differences in unordered categorical variables; and Kruskal-Wallis rank–sum test, to analyze differences in ordinal categorical variables. Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated to quantify the susceptibility to ESCC as determined by each SNP.

We established two adjusted models to minimizing the effects of potential confounding variables. One model was adjusted for age and sex, and another was adjusted for smoking, alcohol intake, education status, wealth score (which was calculated based on the ownership of valuable home items using a multiple correspondence analysis), tea consumption and teeth brushing frequency. The dose–effect relationship between categorical variables and ESCC risk was evaluated with a Chi-square test for trend. The Chi-square test was also used to examine the Hardy-Weinberg equilibrium (HWE), and $P > 0.05$ was considered to indicate equilibrium(37).

Univariate logistic regression was used to assess the association between genotype distribution and ESCC risk in codominant, dominant, recessive and over-dominant models, with the ancestral allele as reference. The Akaike information criteria score was used to select the fittest model for every SNP(38). Inherited model classification and calculation were conducted using the R package "SNPassoc" (https://cran.r-project.org, package = SNPassoc)(39).

Multivariable logistic regression models were used to calculate the statistical significance of SNPs and smoking as indicators of the ESCC risk. We included the SNP–smoking product term as a measurement of interaction on a multiplicative scale, and adjusted for age, sex, alcohol consumption, wealth score, tea consumption and teeth brushing frequency. The significance of the interaction term was tested by the likelihood ratio test through comparison of two models. Models as follows:

$$y_0 = \beta_{00} + \beta_{01} \cdot smoking + \beta_{02} \cdot SNP + \varepsilon_0$$

$$y_1 = \beta_{10} + \beta_{11} \cdot smoking + \beta_{12} \cdot SNP + \beta_{13} \cdot smoking \cdot SNP + \varepsilon_1$$

Moreover, as recommended by Andersson et al. and Knol et al.(40, 41), the relative excess risk of interaction (RERI) and the synergy index (S) were also used to evaluate interaction on an additive scale. RERI was calculated with the R package epiR (https://cran.r-project.org, package = epiR).

For individuals, Genetic Risk Score (GRS) (42, 43)was constructed by the sum of effect allele. The formula is described below:

$$\mathrm{GRS}_i = \sum_{k=1}^{n} G_{i,k},$$

where GRS(i,j) stands for the GRS value for the i_{th} individual, n_j for the numbers of SNPs in ESCC risk, and G(i,k) for the dosage in k_{th} SNP for i_{th} individual.

All statistical analyses were performed on the R software (Version 3.6.2; https://cran.r-project.org/). $P < 0.05$ was considered to indicate statistical significance in all the tests.

The reporting of this study conforms to STROBE guidelines.(44)

Results

Characteristics of study participants

Fifty-two cases and three controls for whom blood samples were not available were excluded from the analysis. Finally, 1030 cases and 1783 controls were used for further analyses. We also removed 26 SNPs that exhibited Hardy-Weinberg Disequilibrium, had a low minor allele frequency, or had greater than 10% missingness across individuals. Thus, 75 SNPs were finally used in the analyses.

The demographic characteristics of the ESCC patients and controls are shown in **Table 1**. The mean age of the patients at the time of diagnosis was 67.00 ± 8.61 years, and the mean age of the controls was 66.30 ± 8.81 years. There was no significant difference between the two groups with regard to sex and marriage status. This was in line with our expectation because the cases and controls were matched for age and sex. Notably, the ESCC patients were significantly to be exposed to smoking and alcohol drinking than the controls. The level of education, frequency of teeth brushing, and wealth score were significantly lower in the ESCC patients than in the controls ($P < 0.001$).

Table 1. Basic characteristics of esophageal squamous cell carcinoma (ESCC) cases and controls

	Case (n = 1030)	Control (n = 1783)	*P* values
Age (Mean ± SD)	67.01±8.61	66.3±8.81	**0.038**
Age group			0.065
40–49	29(2.8)	73(4.1)	
50–59	170(16.5)	326(18.3)	
60–69	430(41.7)	710(39.8)	
70–79	325(31.6)	577(32.4)	
80–85	76(7.4)	97(5.4)	
Sex (%)			0.291
Male	687(66.7)	1225(68.7)	
Female	343(33.3)	558(31.3)	
Smoke intake (%)			**<0.001**
Never	422(41.0)	795(44.6)	
Quitted	48(4.7)	140(7.9)	
Smoking	512(49.7)	821(46.0)	
Missing	48(4.7)	27(1.5)	
Ranked Pack-Years			**0.001**
0	480(46.6)	826(46.3)	
1~30	115(11.2)	283(15.9)	
≥30	435(42.2)	674(37.8)	
Alcohol intake (%)			**<0.001**
Never	458(44.5)	1030(57.8)	
Quitted	26(2.5)	68(3.8)	
Drinking	496(48.2)	656(36.8)	
Missing	50(4.9)	29(1.6)	
Education level (%)			**<0.001**
Illiteracy	376(36.5)	491(27.5)	
Primary school	383(37.2)	678(38.0)	
Junior school	207(20.1)	468(26.2)	
High school and above	64(6.2)	146(8.2)	
Brush times (%)			**<0.001**
<2	817(79.3)	1156(64.8)	
≥2	178(17.3)	604(33.9)	
Missing	35(3.4)	23(1.3)	
Marriage (%)			0.104
Unmarried	38(3.7)	62(3.5)	
Married	786(76.3)	1420(79.6)	
Divorce/Widow	206(20.0)	301(16.9)	
Wealth Grade (%)			**<0.001**

Grade 1	297(28.8)	355(19.9)	
Grade 2	178(17.3)	325(18.2)	
Grade 3	238(23.1)	386(21.6)	
Grade 4	190(18.4)	387(21.7)	
Grade 5	127(12.3)	330(18.5)	
Hot Tea Consumption (%)			
No	668 (68.1)	1280 (72.9)	**<0.001**
Yes	313 (31.9)	475 (27.1)	

SD = standard deviation.

[1] *P* values were derived using chi-square test for categorical variables and anova test for continuous variables.

Dose-effect of tobacco consumption to risk of ESCC

Table 2 presents the relationship between each of the smoking measurement indices we collected and ESCC risk. We established an age and sex-adjusted model and a fully multivariable-adjusted model. Smoking behavior was found to increase the risk of ESCC by 48% (aOR = 1.48, 95% CI = 1.17–1.88) in smokers, in comparison to non-smokers. Further, smoking pack-years had a dose-effect relationship with ESCC risk (Adjusted P for trend = 0.008). Compared with non-smokers, individuals who consume more than 30 pack-years have a 47% higher risk of ESCC (aOR = 1.47, 95% CI = 1.15–1.8). Furthermore, compared with non-smokers, deep inhalation increases the risk of ESCC by 55% (OR = 1.55, 95% CI = 1.21–1.99, according to the fully adjusted model).

Table 2. Odds ratios (ORs) and 95% confidence intervals (CIs) of ESCC associated with smoking behavior

Variables	Case (%)	Control (%)	cOR (95%CI)	aOR (95%CI)
Smoke (%)				
Never	422(41.0)	795(44.6)	1	1
Quitted	48(4.7)	140(7.9)	0.64 (0.45-0.91)	0.79 (0.53-1.17)
Smoking	512(49.7)	821(46.0)	1.17 (0.99-1.37)	1.48 (1.17-1.88)
Ranked Pack-Years				
0	480(46.6)	826(46.3)	1	1
1~30	115(11.2)	283(15.9)	0.75 (0.58-0.95)	0.90 (0.67-1.21)
≥30	435(42.2)	674(37.8)	1.19 (1.01-1.41)	1.47 (1.15-1.87)
P for Trend			<0.001	0.008
Inhalation Smoke				
Never smoker	442(43.1)	795(45.3)	1	1
No	150(15.3)	339(19.3)	0.83 (0.66-1.04)	1.07 (0.81-1.43)
Yes	408(41.6)	620(35.3)	1.24 (1.04-1.47)	1.55 (1.21-1.99)

cOR: crude OR; aOR: OR adjusted by age (continuous), sex, alcohol intake, education, wealth grade, teeth brush frequency and hot tea consumption

Interaction of SNPs and smoking on ESCC

We first evaluated the association between every SNP and ESCC risk using multivariable logistic regression with a codominant model and found that 28 of the 75 SNPs were significantly associated with ESCC risk. All SNPs was corrected by Bonferroni test. Next, the model with the best inheritance from among the codominant, dominant, recessive and over-dominant models of these SNPs was identified. The description of 78 SNPs was shown in **Supplementary Figure 1**. Each of the 78 SNPs were entered into the most significant model among the four models for analyzing its relationship with ESCC. The detail of those genotype and other findings are shown in **Table 3**. The intake form of tobacco was also important, the time exposure had more contribution than dose in ESCC.

Table 3. SNPs Significant (p <0.05) in Codominant Model Analyses for the Association with ESCC (1083 Case Patients and 1786 Controls)

SNP	Gene (or the Nearest)	A/B Alleles	ORab	ORbb	P-Value
rs11066015	ACAD10	G/A	1.21(1.03-1.43)	0.52(0.34-0.8)	<0.001
rs3819197	ADH1A	T/C	0.99(0.84-1.16)	1.85(1.33-2.58)	0.001
rs1159918	ADH1B	G/T	1.03(0.87-1.22)	2.21(1.52-3.21)	<0.001
rs17033	ADH1B	A/G	1.06(0.87-1.3)	2.44(1.27-4.67)	0.023
rs1042026	ADH1B	G/A	0.99(0.84-1.17)	1.96(1.38-2.78)	0.001
rs3805322	ADH4	A/G	0.74(0.61-0.89)	0.87(0.7-1.08)	0.005
rs10008281	ADH6	A/C	0.95(0.81-1.12)	1.45(1.08-1.94)	0.024
rs1893883	ADH6	C/G	1.03(0.86-1.23)	2.04(1.25-3.32)	0.017
rs671	ALDH2	G/A	1.18(1-1.39)	0.52(0.34-0.8)	<0.001
rs3782886	BRAP	A/G	1.16(0.99-1.37)	0.5(0.33-0.76)	<0.001
rs10931936	CASP8	C/T	1.39(1.18-1.64)	1.81(1.38-2.37)	<0.001
rs3769823	CASP8	C/T	1.37(1.16-1.61)	1.74(1.32-2.3)	<0.001
rs10201587	FLACC1	G/A	1.34(1.14-1.58)	1.6(1.21-2.1)	<0.001
rs13016963	FLACC1	G/A	1.39(1.18-1.64)	1.81(1.38-2.36)	<0.001
rs9288318	FLACC1	A/C	1.33(1.13-1.56)	1.71(1.3-2.24)	<0.001
rs4785204	HEATR3	C/T	1.16(0.99-1.36)	1.42(1.05-1.93)	0.031
rs11066280	HECTD4	T/A	1.11(0.94-1.3)	0.59(0.41-0.86)	0.003
rs2074356	HECTD4	C/T	1.14(0.96-1.36)	0.46(0.24-0.88)	0.009
rs2051428	LOC100507053	C/T	1.01(0.86-1.19)	1.42(1.07-1.87)	0.046
rs1442490	LOC102723576	T/C	0.87(0.73-1.03)	1.26(1-1.58)	0.003
rs994771	LOC102723576	T/C	0.83(0.71-0.98)	1.08(0.83-1.4)	0.043
rs11065783	LOC105369980	A/G	1.2(1.02-1.41)	0.88(0.66-1.17)	0.025
rs11187842	PLCE1	C/T	1.47(1.24-1.74)	1.71(1.08-2.72)	<0.001
rs11187870	PLCE1	G/C	1.37(1.16-1.62)	1.8(1.15-2.83)	<0.001
rs3781264	PLCE1	T/C	1.42(1.19-1.68)	1.98(1.27-3.1)	<0.001
rs3765524	PLCE1	C/T	1.49(1.27-1.76)	2.35(1.66-3.31)	<0.001
rs753724	PLCE1	G/T	1.45(1.22-1.72)	1.77(1.13-2.77)	<0.001
rs25487	XRCC1	G/A	0.97(0.83-1.15)	0.57(0.4-0.81)	0.005

We finally found three of these SNPs showed significant association on multiplicative scale, they were ADH1B rs1042026 (recessive model), HECTD4 rs11066280 (over-dominant model) and CASP8 rs3769823 (codominant model). For rs1042026, genotype GG/GA was identified in 2667 (95.14%) samples, while genotype AA was identified in 136 (4.86%) samples. For rs11066280, count for genotype TT/TA was 2637 (94.34%) samples, while AA were identified in 158 (5.66%). And for rs3769823, genotype CC, CT, TT counts for 1418(51.26%) ,1112 (40.20%) and 236 (8.53%).

Table 4 presents the ORs and 95% CI values for rs1042026 of genotype AA with moderate smoking exposure (OR = 2.18, 95% CI = 1.08-14.44) and heavy smoking exposure (OR = 4.25, 95% CI = 2.11–8.98), respectively (P value for trend = 0.001). rs11066280 of genotype TA with moderate smoking exposure (OR = 1.43, 95% CI = 1.01-2.02) and heavy smoking exposure (OR = 1.80, 95% CI = 1.28–2.53), respectively (P value for trend = 0.012). Moreover, rs3769823 CC, CT and TT with heavy smoking exposure showed more than 49% (OR = 1.49, 95% CI = 1.16–1.92) ,60% (OR = 1.60, 95% CI = 1.21–2.13) and 259% (OR = 3.59, 95% CI = 2.14–6.10) excessive risk on ESCC.

Interaction between SNP genotype and cumulative smoking exposure (pack-years) for ESCC risk.

Packs	Cases	Controls	Crude OR (95%CI)	P for trend	Adjusted OR (95%CI)	P for trend
5						
0	402	758	1		1	
<30	204	455	0.85 (0.69-1.04)		0.90 (0.68-1.19)	
≥30	300	473	1.20 (0.99-1.44)	0.005	1.28 (0.98-1.67)	0.004
0	17	33	0.97 (0.52-1.74)		0.86 (0.45-1.59)	
<30	20	18	2.10 (1.09-4.04)		2.18 (1.08-4.44)	
≥30	29	13	4.21 (2.21-8.45)	0.004	4.25 (2.11-8.98)	0.001
0						
0	262	486	1		1	
<30	119	297	0.74 (0.57-0.96)		0.73 (0.52-1.00)	
≥30	197	302	1.21 (0.96-1.53)	0.004	1.16 (0.85-1.58)	0.001
0	160	305	0.97 (0.76-1.24)		1.07 (0.83-1.38)	
<30	107	174	1.14 (0.86-1.51)		1.43 (1.01-2.02)	
≥30	129	176	1.36 (1.03-1.78)	0.02	1.80 (1.28-2.53)	0.012
3						
0	185	421	1		1	
<30	107	247	1.07 (0.75-1.52)		0.99 (0.74-1.31)	
≥30	150	266	1.39 (0.99-1.95)	0.004	1.28 (0.98-1.67)	0.003
0	194	296	1.53 (1.18-1.98)		1.49 (1.16-1.92)	
<30	96	181	1.33 (0.93-1.91)		1.21 (0.89-1.63)	
≥30	131	186	1.73 (1.22-2.47)	0.001	1.60 (1.21-2.13)	0.001
0	38	60	1.38 (0.86-2.17)		1.44 (0.92-2.23)	
<30	21	40	1.14 (0.61-2.08)		1.19 (0.67-2.06)	
≥30	41	26	4.29 (2.42-7.74)	0.002	3.59 (2.14-6.10)	0.002

On additive scale, the RERI and S for rs1042026 AA was consistent with multiplicative scale (RERI = 3.11, S = 23.21), as the same as rs11066280 TA (RERI = 0.57 95%CI = 0.36-0.71, S = 3.48). However, for rs3769823, the RERI and S for genotype CT showed none (RERI = -0.17,95%CI = -0.35-0.12) and weak (S = 0.78) association.

Genetic Risk Score (GRS)

We calculated the Genetic Risk Score (GRS) to quantify the genetic predisposition to ESCC risk. The GRS varied from 0 to 5. We further defined the GRS as low (0-1), moderate (2-3) and high (4-5) GRS score, the moderate and high GRS could increase 1.29 (1.10-1.50) and 1.87(1.25-2.80) times of risk on ESCC. As shown on **Figure 2,** when stratified by whether smoke, in non-smoker the genetic predisposition showed unsignificant effect on ESCC (OR = 1.15, 95%CI = 0.98-1.58; OR = 0.93, 95%CI = 0.45-1.79); strong association between ESCC and genetic mutation were found in smokers, it could increase 34% and 192% of cancer risk (OR = 1.34, 95%CI = 1.01-1.54; OR = 2.92, 95%CI = 1.69-5.11).

Discussion

To our knowledge, this is the first study to investigate the interaction effect of SNP and smoking on the risk of ESCC in a large Chinese population. In this study, we analyzed 1030 cases and 1783 controls to explore the interaction effect between smoking and candidate SNPs associated with ESCC risk. The aim of the present study was to investigate the interaction effect of smoking and SNPs, significant loci were identified in three different genes.

The analyses in this study were conducted according to the recommendations of Andersson et al. and Knol et al., who showed that biological interaction should be evaluated on both the multiplicative and additive scales. In this study, there showed positive interaction effect on both the multiplicative scale and additive scale in rs1042026 and rs11066280, the result from the two scales is almost consistent, only for CASP8 rs3769823 CT, the OR showed 1.60 times of interaction risk compared with reference group, while RERI showed unsignificant difference on additive scale (RERI = -0.17). This inconsistence showed statistical and biological interaction have complex relationship. So, we retrieved publication for further evidence, we found that CASP8 encodes a member of the cysteine-aspartic acid protease (caspase) family and involved in the programmed cell death induced by Fas and various apoptotic stimuli, which implies that it has association between tumor progress. Moreover, we found that Jong et.al(45) analyzed rs3769823 was significantly enriched in tumor necrosis factor pathway and apoptosis pathway in lung and smoking exposure research. Therefore, we thought the rs3769823 have interaction effect between ESCC and smoking, but more research was needed to prove this point of view.

Rs1042026 was identified as significant loci conducted in this study. This SNPs was located on the ADH1B gene, which participates in alcohol-metabolizing pathways and contributes significantly to elevating the risk of ESCC(46, 47). In our previous study, we found that genes associated with alcohol metabolism, such as ADH1B and ALDH2, interact with alcohol consumption to significantly increase the risk of developing ESCC(28, 48). Further, a meta-analysis reported that alcohol drinking combined with smoking was associated with twice the risk of ESCC than exposure to only alcohol or only smoking. With regard to the underlying mechanism, it has been shown that alcohol damages cellular DNA by decreasing metabolic activity within the cell and, therefore, reduces detoxification and promotes oxidation. In this

study, although we have adjusted alcohol as covariable, the ADH1B rs1042026 was still showed strong association in interaction effect. Researchers have found that smoking and alcohol intake together have a synergistic effect on the risk of ESCC (49), therefore, we thought alcohol drinkers tend to also smoke, and alcohol contributes more than smoking to increasing the risk of ESCC. Considering that alcohol intake has a correlation with both smoking and ESCC, alcohol intake may be a confounding factor. Therefore, we performed stratified analysis to account for the influence of alcohol intake, we found effect in alcohol-intake group. Thus, alcohol seems to be a major risk factor for this disease and smoking alone seems to carry a limited risk of ESCC. Nonetheless, the findings indicate that the alcohol-related SNPs have an interactive effect with smoking.

GRS was a quantitative variable by which people can learn about their risk of developing a disease, based on the total number of changes related to the disease. The GRS was calculated by their all-selected variants to combine all potential and their contribution in specific disease. In this study, we summarized three of these SNPs to establish the GRS, we found that there was no significant effect in non-smokers, although individuals' genetic risk varied from low to high level, it implied that these SNPs do not rise excessive risk on ESCC when there was no tobacco exposure, while in smokers with highest genetic risk, interaction effect was found to increase nearly two times of risk on ESCC. Taking screening of these SNPs in smokers could seek out the potential population and give prevention measure.

The main limitation of this study is that we could not give definitive conclusion to those SNPs, because we can only find few related research or function introduction in literature database. Moreover, the ESCC risk in moderate smokers was not significant, and moderate smokers only accounted for a small percentage of the total patients with a mutant genotype (resulting in a wide 95% CI).

In summary, the findings of this large population-based case-control study on the potential interactions between SNPs and cigarettes consumption exposure and their effect on ESCC risk indicate that smoking-related factors have an interactive effect with smoking and carry a high risk on ESCC. We believe that the findings provide important insight into risk stratification of the population and disease prevention.

Declarations

Acknowledgements

We would like to thank the interviewers and technicians of Fudan-Taizhou Institute of Health Sciences for their invaluable contribution to the data collection and sample preparation, the staffs of Taixing Center for Disease Control and Prevention for helping in the field work, and the staffs of Taixing People's Hospital for the assistance of sample collection.

Authors' contributions

RZ designed the study, made substantial contributions to conception, and contributed to writing the manuscript. HY participated in writing the manuscript and data analysis. ZL, XL and HC participated in the design of the study and in critical revision of the manuscript draft. XY, TZ, ML, and LJ contributed to design of the study, conceived data recording and analysis. XC, CS conceived the study, performed experiments, and essentially wrote the manuscript. All authors provided comments to the draft manuscript and read and approved the final manuscript.

Funding

This study was supported by the National Key Research and Development program of China (grant number: 2017YFC0907000, 2019YFC1315804, 2017YFC0907500, 2016YFC0901403), the National Natural Science Foundation of China (grant numbers: 91846302, 81502870); the Innovation Grant from Science and Technology Commission of Shanghai Municipality, China (grant number: 20ZR1405600)； Three-Year Action Plan for Strengthening Public Health System in Shanghai (grant number: GWV-10.2-YQ32)；the Key Basic Research Grants from Science and Technology Commission of Shanghai Municipality, China (grant number: 16JC1400501), Shanghai Municipal Science and Technology Major Project (grant number: 2017SHZDZX01), and the International Science and Technology Cooperation Program of China (grant number: 2015DFE32790).

Availability of data and materials

All data that support the findings of this study are available from the corresponding authors for a reasonable request.

Ethics approval and consent to participate

Ethical approval was sought from Chinese Ethic committee in this study.

Consent for publication

Not applicable.

Competing interests

The authors declare no potential conflicts of interest.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. Ca-Cancer J Clin. 2019;69(1):7–34.
2. Henry MA, Lerco MM, Ribeiro PW, Rodrigues MA. Epidemiological features of esophageal cancer. Squamous cell carcinoma versus adenocarcinoma. Acta Cir Bras. 2014;29(6):389–93.
3. Abnet CC, Arnold M, Wei WQ. Epidemiology of Esophageal Squamous Cell Carcinoma. Gastroenterology. 2018;154(2):360–73.
4. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. Ca-Cancer J Clin. 2016;66(2):115–32.
5. Fan J, Liu Z, Mao X, Tong X, Zhang T, Suo C, et al. Global trends in the incidence and mortality of esophageal cancer from 1990 to 2017. Cancer Med. 2020:e03338.
6. Wang Q-L, Xie S-H, Li W-T, Lagergren J. Smoking Cessation and Risk of Esophageal Cancer by Histological Type: Systematic Review and Meta-analysis. Journal of the National Cancer Institute. 2017;109(12).
7. Najafi F. Tobacco Smoking and Alcohol Drinking: Two Clinically Significant Risk Factors for Esophageal Squamous Cell Carcinoma. Gastroenterology. 2019;157(3):897.
8. Meng XY, Zhu ST, Zhou QZ, Li P, Wang YJ, Zhang ST. Promoter methylation regulates cigarette smoke-stimulated cyclooxygenase-2 expression in esophageal squamous cell carcinoma. J Dig Dis. 2012;13(4):208–13.
9. Mao N, Nie S, Hong B, Li C, Shen X, Xiong T. Association between alcohol dehydrogenase-2 gene polymorphism and esophageal cancer risk: a meta-analysis. World J Surg Oncol. 2016;14(1):191.
10. Yang X, Chen X, Zhuang M, Yuan Z, Nie S, Lu M, et al. Smoking and alcohol drinking in relation to the risk of esophageal squamous cell carcinoma: A population-based case-control study in China. Scientific Reports. 2017;7(1).
11. Freedman ND, Abnet CC, Leitzmann MF, Mouw T, Subar AF, Hollenbeck AR, et al. A prospective study of tobacco, alcohol, and the risk of esophageal and gastric cancer subtypes. Am J Epidemiol. 2007;165(12):1424–33.
12. Siddiqi K, Shah S, Abbas SM, Vidyasagaran A, Jawad M, Dogar O, et al. Global burden of disease due to smokeless tobacco consumption in adults: analysis of data from 113 countries. BMC Med. 2015;13:194.
13. Sewram V, Sitas F, O'Connell D, Myers J. Tobacco and alcohol as risk factors for oesophageal cancer in a high incidence area in South Africa. Cancer Epidemiol. 2016;41:113–21.
14. Cancer Genome Atlas Research N, Analysis Working Group: Asan U, Agency BCC, Brigham, Women's H, Broad I, et al. Integrated genomic characterization of oesophageal carcinoma. Nature. 2017;541(7636):169–75.
15. Qin HD, Liao XY, Chen YB, Huang SY, Xue WQ, Li FF, et al. Genomic Characterization of Esophageal Squamous Cell Carcinoma Reveals Critical Genes Underlying Tumorigenesis and Poor Prognosis. Am J Hum Genet. 2016;98(4):709–27.

16. Lin DC, Wang MR, Koeffler HP. Genomic and Epigenomic Aberrations in Esophageal Squamous Cell Carcinoma and Implications for Patients. Gastroenterology. 2018;154(2):374–89.
17. Song X, Li WQ, Hu N, Zhao XK, Wang Z, Hyland PL, et al. GWAS follow-up study of esophageal squamous cell carcinoma identifies potential genetic loci associated with family history of upper gastrointestinal cancer. Sci Rep. 2017;7(1):4642.
18. Qing T, Zhu S, Suo C, Zhang L, Zheng Y, Shi L. Somatic mutations in ZFHX4 gene are associated with poor overall survival of Chinese esophageal squamous cell carcinoma patients. Sci Rep. 2017;7(1):4951.
19. Zhang Q, Zhao X, Zhang C, Wang W, Li F, Liu D, et al. Overexpressed PKMYT1 promotes tumor progression and associates with poor survival in esophageal squamous cell carcinoma. Cancer Manag Res. 2019;11:7813–24.
20. Sun MZ, Ju HX, Zhou ZW, Jin H, Zhu R. Single nucleotide polymorphisms of DNA mismatch repair genes MSH2 and MLH1 confer susceptibility to esophageal cancer. Int J Clin Exp Med. 2014;7(8):2329–33.
21. Wu M, Chang SC, Kampman E, Yang J, Wang XS, Gu XP, et al. Single nucleotide polymorphisms of ADH1B, ADH1C and ALDH2 genes and esophageal cancer: a population-based case-control study in China. Int J Cancer. 2013;132(8):1868–77.
22. Findlay JM, Middleton MR, Tomlinson I. A systematic review and meta-analysis of somatic and germline DNA sequence biomarkers of esophageal cancer survival, therapy response and stage. Ann Oncol. 2015;26(4):624–44.
23. Li S, Deng Y, You J-P, Chen Z-P, Peng Q-L, Huang X-M, et al. XRCC1 Arg399Gln, Arg194Trp, and Arg280His polymorphisms in esophageal cancer risk: a meta-analysis. Dig Dis Sci. 2013;58(7):1880–90.
24. Gong F-F, Lu S-S, Hu C-Y, Qian Z-Z, Feng F, Wu Y-L, et al. Cytochrome P450 1A1 (CYP1A1) polymorphism and susceptibility to esophageal cancer: an updated meta-analysis of 27 studies. Tumour Biol. 2014;35(10):10351–61.
25. Wang B, Wang D, Zhang D, Li A, Liu D, Liu H, et al. Pro variant of TP53 Arg72Pro contributes to esophageal squamous cell carcinoma risk: evidence from a meta-analysis. Eur J Cancer Prev. 2010;19(4):299–307.
26. Wang Y, Guo W, He Y, Chen Z, Wen D, Zhang X, et al. Association of MTHFR C677T and SHMT(1) C1420T with susceptibility to ESCC and GCA in a high incident region of Northern China. Cancer Causes Control. 2007;18(2):143–52.
27. Luo Q, Zeng J, Li W, Lin L, Zhou X, Tian X, et al. Interaction of MTHFR gene with smoking and alcohol use and haplotype combination susceptibility to psoriasis in Chinese population. Immunol Res. 2018;66(4):543–7.
28. Suo C, Yang Y, Yuan Z, Zhang T, Yang X, Qing T, et al. Alcohol Intake Interacts with Functional Genetic Polymorphisms of Aldehyde Dehydrogenase (ALDH2) and Alcohol Dehydrogenase (ADH) to Increase Esophageal Squamous Cell Cancer Risk. J Thorac Oncol. 2019;14(4):712–25.

29. Ekheden I, Yang X, Chen H, Chen X, Yuan Z, Jin L, et al. Associations Between Gastric Atrophy and Its Interaction With Poor Oral Health and the Risk for Esophageal Squamous Cell Carcinoma in a High-Risk Region of China: A Population-Based Case-Control Study. Am J Epidemiol. 2020;189(9):931–41.
30. Suo C, Qing T, Liu Z, Yang X, Yuan Z, Yang Y-J, et al. Differential Cumulative Risk of Genetic Polymorphisms in Familial and Nonfamilial Esophageal Squamous Cell Carcinoma. Cancer Epidemiology Biomarkers & Prevention. 2019;28(12):2014–21.
31. Chen T, Cheng H, Chen X, Yuan Z, Yang X, Zhuang M, et al. Family history of esophageal cancer increases the risk of esophageal squamous cell carcinoma. Scientific Reports. 2015;5(1).
32. Yang X, Ni Y, Yuan Z, Chen H, Plymoth A, Jin L, et al. Very hot tea drinking increases esophageal squamous cell carcinoma risk in a high-risk area of China: a population-based case-control study. Clin Epidemiol. 2018;10:1307–20.
33. Lee Y-H, Shin M-H, Kweon S-S, Choi J-S, Rhee J-A, Ahn H-R, et al. Cumulative smoking exposure, duration of smoking cessation, and peripheral arterial disease in middle-aged and older Korean men. BMC Public Health. 2011;11:94.
34. Global H. Study Quantifies Smoking Strong Link to Peripheral Artery Disease https://www.hmpgloballearningnetwork.com/site/jic/news/study-quantifies-smokings-strong-link-peripheral-artery-disease2019 [
35. Institute NC. National Cancer Institute definition of pack year http://www.cancer.gov/dictionary?CdrID=306510 [
36. Hopkinsmedicine. Smoking and Cardiovascular Disease https://2021 [
37. Liu H, Hu Y. [Hardy-Weinberg equilibrium in genetic epidemiology]. Zhong Nan Da Xue Xue Bao Yi Xue Ban. 2010;35(1):90–3.
38. Portet S. A primer on model selection using the Akaike Information Criterion. Infect Dis Model. 2020;5:111–28.
39. González JR, Armengol L, Solé X, Guinó E, Mercader JM, Estivill X, et al. SNPassoc: an R package to perform whole genome association studies. Bioinformatics. 2007;23(5):644–5.
40. Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. Int J Epidemiol. 2012;41(2):514–20.
41. Andersson T, Alfredsson L, Källberg H, Zdravkovic S, Ahlbom A. Calculating measures of biological interaction. Eur J Epidemiol. 2005;20(7):575–9.
42. Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc. 2020;15(9):2759–72.
43. Rutten-Jacobs LC, Larsson SC, Malik R, Rannikmae K, consortium M, International Stroke Genetics C, et al. Genetic risk, incident stroke, and the benefits of adhering to a healthy lifestyle: cohort study of 306 473 UK Biobank participants. BMJ. 2018;363:k4168.
44. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting

observational studies. Ann Intern Med. 2007;147(8):573–7.

45. de Jong K, Vonk JM, Imboden M, Lahousse L, Hofman A, Brusselle GG, et al. Genes and pathways underlying susceptibility to impaired lung function in the context of environmental tobacco smoke exposure. Respir Res. 2017;18(1):142.
46. Preuss UW, Ridinger M, Rujescu D, Samochowiec J, Fehr C, Wurst FM, et al. Association of ADH4 genetic variants with alcohol dependence risk and related phenotypes: results from a larger multicenter association study. Addict Biol. 2011;16(2):323–33.
47. Birley AJ, James MR, Dickson PA, Montgomery GW, Heath AC, Martin NG, et al. ADH single nucleotide polymorphism associations with alcohol metabolism in vivo. Human molecular genetics. 2009;18(8):1533–42.
48. Chen X-P, Xu D-F, Xu W-H, Ma Z-C, Yao J, Fu S-M. Association Studies of CYP1A1 Exon7 Polymorphism and -GSTM1 Interaction with Esophageal Cancer Risk: a Meta-Analysis in the Chinese Population. Clin Lab. 2016;62(9):1795–802.
49. Yaegashi Y, Onoda T, Morioka S, Hashimoto T, Takeshita T, Sakata K, et al. Joint effects of smoking and alcohol drinking on esophageal cancer mortality in Japanese men: findings from the Japan collaborative cohort study. Asian Pac J Cancer Prev. 2014;15(2):1023–9.

Figures

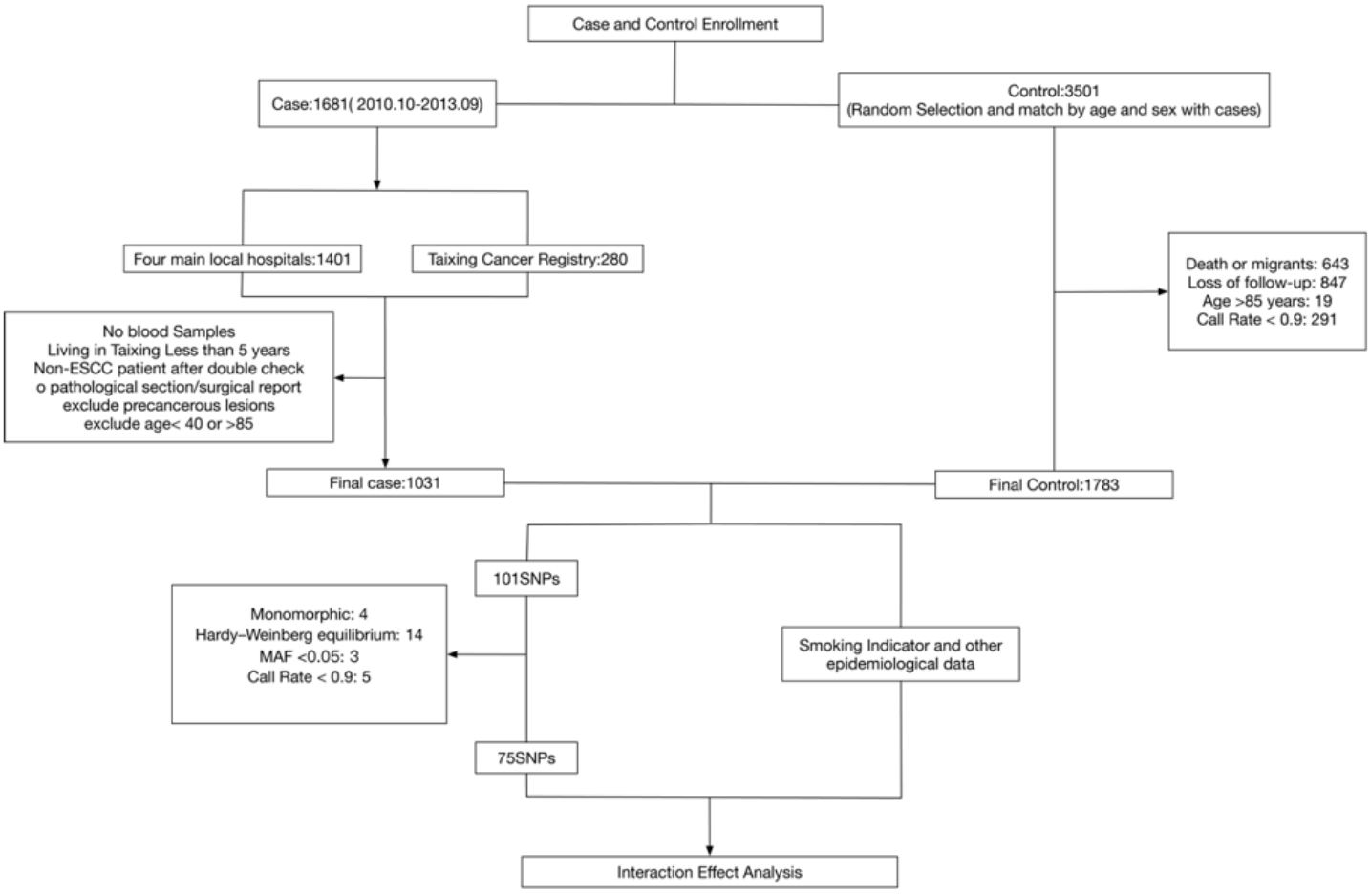

Figure 1

The flow chart for the recruitment of esophageal squamous cell carcinoma cases and controls in a population-based case-control study conducted in Taixing, China.

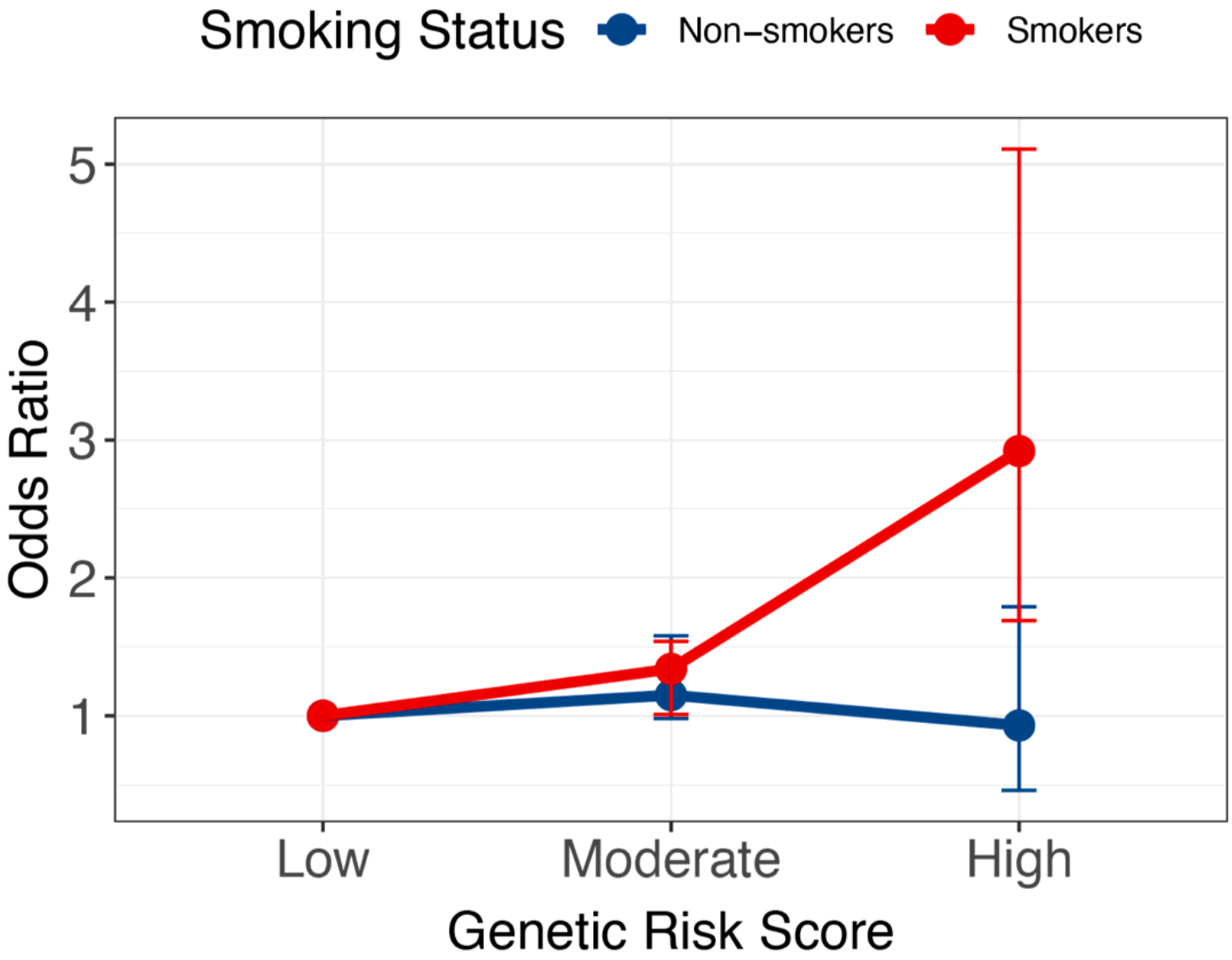

Figure 2

The risk stratification of GRS performed on tobacco pack-years, estimated by esophageal squamous cell carcinoma risk.