

Additional file 1: The definition of comorbidity. Table S1. Test for multicollinearity among the independent variables entered into the multivariable regression. **Table S2.** Comparative analysis between included and excluded subjects. **Table S3.** Risk of oral and pharyngeal cancers based on regarding the variables. **Table S4.** Sensitivity analysis in non-smokers (N=97,037). **Table S5.** Sensitivity analysis in former smokers (N=15,590). **Table S6.** Sensitivity analysis in current smokers (N=38,147). **Table S7.** Sensitivity analysis in non-drinkers (N=78,974). **Table S8.** Sensitivity analysis in drinkers (N=72,224). **Table S9.** Subgroup analysis of occurrence of oral cavity and oral cancers based on regarding the variables. **Table S10.** Subgroup analysis regarding the occurrence of oral cavity and oral cancer. **Table S11.** Subgroup analysis regarding the occurrence of salivary gland cancer. **Table S12.** Subgroup analysis regarding the occurrence of tonsil and pharynx cancer.

The definition of comorbidity

Information on smoking habits and alcohol consumption were obtained by questionnaire.

Body mass index was defined as the participant's weight in kilograms divided by the square of the participant's height in meters. Regular physical exercise was considered to be strenuous physical activity performed for at least 20 min more than once per week [1].

Smoking status was categorized into none, former smoker and current smoker regardless of the amount of smoking. Economic status was dichotomized at the bottom 10%. Hypertension was defined using International Statistical Classification of Diseases, 10th Revision (ICD-10) codes I10-I11 and the prescription of an antihypertensive agent with at least one claim per year. Diabetes mellitus was defined as main or secondary diagnosis of diabetes mellitus (ICD-10 codes E10-E14). Criteria for this diagnosis included at least one claim per year for both visiting an outpatient clinic and admission accompanied by prescription records for any hypoglycemic agents. Alternatively, at least one fasting plasma glucose ≥ 126 mg/dL.

Dyslipidemia was defined using ICD-10 code E78 and the prescription of a lipid-lowering agent including statins with at least one claim per year. Renal disease was identified by ICD-10 codes N18.1-N18.5 and N18.9 with at least one claim per year. Malignancy was defined by ICD-10 codes C00-D48 with at least one claim per year. Body mass index and blood and urine laboratory findings were used with the baseline dataset [2].

References

1. Kim MK, Han K, Joung HN, Baek KH, Song KH, Kwon HS. Cholesterol levels and development of cardiovascular disease in Koreans with type 2 diabetes mellitus and without pre-existing cardiovascular disease. *Cardiovasc Diabetol*. 2019;18(1):139.
2. Park SY, Kim SH, Kang SH, et al. Improved oral hygiene care attenuates the cardiovascular risk of oral health disease: a population-based study from Korea. *Eur Heart J*. 2019;40(14):1138-1145.

Table S1. Test for multicollinearity among the independent variables entered into the multivariable regression for occurrence of oral and pharyngeal cancers as dependent variable.

	Variance inflation factor
Age (years)	1.3129
Male sex	1.6904
Income levels	
Fifth quintile (highest)	Reference
Fourth quintile	1.4264
Third quintile	1.4060
Second quintile	1.3480
First quintile (lowest)	1.4155
Covered by medical aid	1.0256
Regular physical activity	1.0134
Alcohol consumption	1.4325
Body mass index	1.1374
Hypertension	1.7341
Diabetes mellitus	1.5459
Dyslipidemia	1.4799
Smoking status	
Non-smoker	Reference
Former smoker	1.1704
Current smoker	1.4157
Systolic blood pressure	1.7101
Total cholesterol	1.4981
Fasting blood sugar	1.5026
Aspartate aminotransferase	2.6660
Alanine aminotransferase	2.5045
Gamma glutamyl transferase	1.5278
Proteinuria	1.0104
Periodontal disease	1.0548
Frequency of tooth brushings (times/day)	
0–1	Reference
2	1.7348
≥3	1.7992
Dental visits for any reason	1.2556
Professional teeth scaling	1.2686
Number of teeth lost	
0	Reference
1–7	1.0574
8–14	1.0506
≥15	1.0968

Table S2. Comparative analysis between included and excluded subjects.

Characteristics	Included subjects	Excluded subjects	p value	Standardized Difference
Number of subjects	150,774	364,092		
Age (years)	52.2 ± 8.7	55.5 ± 9.8	<.001	0.353
Male sex, % (n)	91,855 (60.9)	187,270 (51.4)	<.001	-0.192
Income levels, % (n)			<.001	0.184
Fifth quintile (highest)	59,339 (39.4)	116,521 (32.0)		
Fourth quintile	30,701 (20.4)	77,243 (21.2)		
Third quintile	20,871 (13.8)	59,576 (16.4)		
Second quintile	18,962 (12.6)	51,427 (14.1)		
First quintile (lowest)	20,665 (13.7)	56,149 (15.4)		
Covered by medical aid	236 (0.2)	3,176 (0.9)		

P value by Student's t-test and Chi-square test.

Table S3. Risk of oral and pharyngeal cancers based on regarding the variables.

Variables	Univariable analysis		Multivariable analysis	
	HR (95% CI)	P value	HR (95% CI)	P value
Age (years)	1.08 (1.08–1.09)	<.001	1.07 (1.07–1.08)	<.001
Male sex	0.55 (0.49–0.61)	<.001	0.65 (0.56–0.76)	<.001
Income levels				
Fifth quintile (highest)	1 (Reference)		1 (Reference)	
Fourth quintile	1.55 (1.31–1.83)	<.001	1.35 (1.14–1.59)	<.001
Third quintile	1.96 (1.65–2.33)	<.001	1.58 (1.32–1.88)	<.001
Second quintile	1.66 (1.38–2.01)	<.001	1.18 (0.97–1.43)	0.096
First quintile (lowest)	1.81 (1.51–2.16)	<.001	1.11 (0.93–1.33)	0.258
Covered by medical aid	5.62 (2.66–11.89)	<.001	2.18 (1.03–4.63)	0.041
Alcohol consumption	0.64 (0.57–0.72)	<.001	1.08 (0.94–1.25)	0.269
Smoking status				
Non-smoker	1 (Reference)		1 (Reference)	
Former smoker	0.54 (0.42–0.68)	<.001	0.82 (0.64–1.06)	0.123
Current smoker	0.62 (0.53–0.72)	<.001	0.92 (0.78–1.10)	0.375
Regular physical activity	1.05 (0.87–1.28)	0.587	0.93 (0.77–1.13)	0.461
Body mass index	1.01 (0.98–1.03)	0.642	1.02 (1.00–1.04)	0.091
SBP (per 10 unit increase)	1.07 (1.03–1.10)	<.001	0.94 (0.91–0.98)	0.005
Comorbidities				
Hypertension	1.49 (1.33–1.67)	<.001	1.15 (0.98–1.34)	0.083
Diabetes mellitus	1.52 (1.28–1.80)	<.001	1.00 (0.80–1.24)	0.969
Dyslipidemia	1.05 (0.90–1.22)	0.569	0.96 (0.80–1.16)	0.682
Laboratory findings				
Total cholesterol (per 10 unit increase)	0.99 (0.97–1.01)	0.180	0.98 (0.96–1.00)	0.018
FBS (per 10 unit increase)	1.03 (1.02–1.04)	<.001	1.02 (1.00–1.04)	0.023
AST (per 10 unit increase)	1.03 (1.02–1.05)	<.001	1.03 (0.99–1.07)	0.096
ALT (per 10 unit increase)	0.98 (0.94–1.01)	0.212	0.98 (0.95–1.02)	0.438
GGT (per 10 unit increase)	0.99 (0.87–1.13)	0.889	1.07 (0.98–1.17)	0.134
Proteinuria	1.08 (0.78–1.49)	0.657	0.98 (0.71–1.37)	0.919
Oral hygiene care				
Frequency of tooth brushings (times/day)				
0–1	1 (Reference)		1 (Reference)	
2	0.80 (0.69–0.93)	0.004	0.94 (0.80–1.09)	0.397
≥3	0.49 (0.42–0.58)	<.001	0.79 (0.66–0.94)	0.007
Dental visits for any reason	0.88 (0.79–0.99)	0.040	1.06 (0.93–1.21)	0.377
Professional teeth scaling	0.65 (0.56–0.76)	<.001	0.91 (0.77–1.08)	0.295
Oral health status				
Periodontal disease	0.76 (0.63–0.92)	0.004	0.80 (0.65–1.03)	0.063
Number of teeth lost				
0	1 (Reference)		1 (Reference)	
1–7	1.47 (1.29–1.68)	<.001	1.29 (1.12–1.47)	<.001
8–14	3.02 (2.23–4.10)	<.001	1.47 (1.08–2.01)	0.014
≥15	4.93 (3.69–6.59)	<.001	1.66 (1.22–2.25)	0.001

Multivariable model was adjusted for age, sex, income level, alcohol consumption, smoking status, regular exercise, body mass index (kg/m²), hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria,

periodontal disease, frequency of tooth brushings, dental visits for any reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

HR, hazard ratio; CI, confidence interval; SBP, systolic blood pressure; FBS, fasting blood sugar; AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, gamma glutamyl transferase.

Table S4. Sensitivity analysis in non-smokers (N=97,037).

Variables	Multivariable analysis	
	HR (95% CI)	p value
Periodontal disease		
No	1 (Reference)	
Yes	0.73 (0.58–1.02)	0.062
Frequency of tooth brushings (times/day)		
0–1	1 (Reference)	
2	0.95 (0.80–1.14)	0.612
≥3	0.79 (0.65–0.96)	0.021
P for trend*	0.012	
Dental visits for any reason		
No	1 (Reference)	
Yes	1.06 (0.91–1.23)	0.469
Professional teeth scaling		
No	1 (Reference)	
Yes	0.82 (0.66–1.01)	0.067
Number of teeth lost		
0	1 (Reference)	
1–7	1.27 (1.08–1.48)	0.003
8–14	1.56 (1.10–2.21)	0.012
≥15	1.61 (1.14–2.29)	0.007
P for trend*	<.001	

Multivariable model was adjusted for age, sex, income level, alcohol consumption, regular exercise, body mass index (kg/m²), hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria, periodontal disease, frequency of tooth brushings, dental visits for any reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

HR, hazard ratio; CI, confidence interval.

*: The trend test for hazard ratios.

Multivariable analysis was performed by using regression methods of Fine and Gray for competing risk data (death is a competing event for oral and pharyngeal cancers).

Table S5. Sensitivity analysis in former smokers (N=15,590).

Variables	Multivariable analysis	
	HR (95% CI)	p value
Periodontal disease		
No	1 (Reference)	
Yes	0.73 (0.36–1.46)	0.370
Frequency of tooth brushings (times/day)		
0–1	1 (Reference)	
2	1.26 (0.66–2.37)	0.484
≥3	1.08 (0.53–2.18)	0.831
P for trend*	0.946	
Dental visits for any reason		
No	1 (Reference)	
Yes	1.25 (0.73–2.12)	0.417
Professional teeth scaling		
No	1 (Reference)	
Yes	1.11 (0.62–2.00)	0.725
Number of teeth lost		
0	1 (Reference)	
1–7	1.63 (0.98–2.73)	0.060
8–14	0.75 (0.10–5.63)	0.780
≥15	2.92 (0.77–11.08)	0.116
P for trend*	0.053	

Multivariable model was adjusted for age, sex, income level, alcohol consumption, regular exercise, body mass index (kg/m²), hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria, periodontal disease, frequency of tooth brushings, dental visits for any reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

HR, hazard ratio; CI, confidence interval.

*: The trend test for hazard ratios.

Multivariable analysis was performed by using regression methods of Fine and Gray for competing risk data (death is a competing event for oral and pharyngeal cancers).

Table S6. Sensitivity analysis in current smokers (N=38,147).

Variables	Multivariable analysis	
	HR (95% CI)	p value
Periodontal disease		
No	1 (Reference)	
Yes	1.02 (0.68–1.52)	0.936
Frequency of tooth brushings (times/day)		
0–1	1 (Reference)	
2	0.79 (0.57–1.10)	0.166
≥3	0.73 (0.50–1.08)	0.111
P for trend*	0.109	
Dental visits for any reason		
No	1 (Reference)	
Yes	1.02 (0.75–1.40)	0.885
Professional teeth scaling		
No	1 (Reference)	
Yes	1.18 (0.82–1.69)	0.377
Number of teeth lost		
0	1 (Reference)	
1–7	1.24 (0.91–1.68)	0.166
8–14	1.36 (0.65–2.85)	0.413
≥15	1.69 (0.83–3.44)	0.151
P for trend*	0.058	

Multivariable model was adjusted for age, sex, income level, alcohol consumption, regular exercise, body mass index (kg/m²), hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria, periodontal disease, frequency of tooth brushings, dental visits for any reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

HR, hazard ratio; CI, confidence interval.

*: The trend test for hazard ratios.

Multivariable analysis was performed by using regression methods of Fine and Gray for competing risk data (death is a competing event for oral and pharyngeal cancers).

Table S7. Sensitivity analysis in non-drinkers (N=78,974).

Variables	Multivariable analysis	
	HR (95% CI)	p value
Periodontal disease		
No	1 (Reference)	
Yes	0.78 (0.60–1.01)	0.056
Frequency of tooth brushings (times/day)		
0–1	1 (Reference)	
2	0.92 (0.75–1.11)	0.377
≥3	0.78 (0.63–0.97)	0.027
P for trend*	0.022	
Dental visits for any reason		
No	1 (Reference)	
Yes	1.05 (0.89–1.24)	0.580
Professional teeth scaling		
No	1 (Reference)	
Yes	0.88 (0.70–1.11)	0.289
Number of teeth lost		
0	1 (Reference)	
1–7	1.35 (1.14–1.60)	<.001
8–14	1.86 (1.32–2.62)	<.001
≥15	1.74 (1.22–2.48)	0.002
P for trend*	<.001	

Multivariable model was adjusted for age, sex, income level, smoking status, regular exercise, body mass index (kg/m²), hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria, periodontal disease, frequency of tooth brushings, dental visits for any reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

HR, hazard ratio; CI, confidence interval.

*: The trend test for hazard ratios.

Multivariable analysis was performed by using regression methods of Fine and Gray for competing risk data (death is a competing event for oral and pharyngeal cancers).

Table S8. Sensitivity analysis in drinkers (N=72,224).

Variables	Multivariable analysis	
	HR (95% CI)	p value
Periodontal disease		
No	1 (Reference)	
Yes	0.82 (0.60–1.12)	0.210
Frequency of tooth brushings (times/day)		
0–1	1 (Reference)	
2	0.98 (0.76–1.26)	0.874
≥3	0.81 (0.61–1.07)	0.134
P for trend*	0.105	
Dental visits for any reason		
No	1 (Reference)	
Yes	1.09 (0.87–1.35)	0.459
Professional teeth scaling		
No	1 (Reference)	
Yes	0.95 (0.73–1.24)	0.712
Number of teeth lost		
0	1 (Reference)	
1–7	1.17 (0.94–1.46)	0.155
8–14	0.69 (0.32–1.48)	0.338
≥15	1.53 (0.83–2.81)	0.169
P for trend*	0.207	

Multivariable model was adjusted for age, sex, income level, smoking status, regular exercise, body mass index (kg/m²), hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria, periodontal disease, frequency of tooth brushings, dental visits for any reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

HR, hazard ratio; CI, confidence interval.

*: The trend test for hazard ratios.

Multivariable analysis was performed by using regression methods of Fine and Gray for competing risk data (death is a competing event for oral and pharyngeal cancers).

Table S9. Subgroup analysis of occurrence of oral cavity and oral cancers (ICD-10 codes C00-06) based on regarding the variables.

Variables	Univariable analysis		Multivariable analysis	
	HR (95% CI)	p value	HR (95% CI)	p value
Age (years)	1.07 (1.07–1.09)	<.001	1.06 (1.05–1.08)	<.001
Male sex	0.55 (0.50–0.62)	<.001	0.64 (0.54–0.77)	<.001
Income levels				
Fifth quintile (highest)	1 (Reference)		1 (Reference)	
Fourth quintile	1.57 (1.29–1.81)	<.001	1.35 (0.98–1.54)	0.174
Third quintile	1.97 (1.63–2.31)	<.001	1.46 (0.98–1.67)	0.189
Second quintile	1.69 (1.34–2.02)	<.001	1.14 (0.93–1.38)	0.178
First quintile (lowest)	1.84 (1.53–2.19)	<.001	1.19 (0.92–1.57)	0.328
Covered by medical aid	5.64 (2.62–11.94)	<.001	2.19 (1.04–4.15)	0.042
Alcohol consumption	0.65 (0.57–0.72)	<.001	1.04 (0.97–1.24)	0.257
Smoking status				
Non-smoker	1 (Reference)		1 (Reference)	
Former smoker	0.53 (0.41–0.66)	<.001	0.89 (0.65–1.05)	0.108
Current smoker	0.64 (0.54–0.71)	<.001	0.97 (0.77–1.11)	0.461
Regular physical activity	1.06 (0.89–1.29)	0.585	0.92 (0.75–1.12)	0.354
Body mass index	1.02 (0.99–1.05)	0.644	1.03 (1.00–1.04)	0.089
SBP (per 10 unit increase)	1.06 (1.04–1.09)	<.001	1.04 (1.03–1.08)	0.004
Comorbidities				
Hypertension	1.48 (1.34–1.68)	<.001	1.14 (0.98–1.31)	0.175
Diabetes mellitus	1.54 (1.27–1.79)	<.001	1.00 (0.78–1.75)	0.955
Dyslipidemia	1.06 (0.92–1.24)	0.576	1.04 (0.80–1.16)	0.897
Laboratory findings				
Total cholesterol (per 10 unit increase)	0.99 (0.98–1.01)	0.182	0.99 (0.96–1.00)	0.089
FBS (per 10 unit increase)	1.03 (1.02–1.04)	<.001	1.02 (1.00–1.04)	0.024
AST (per 10 unit increase)	1.03 (1.02–1.06)	<.001	1.03 (0.99–1.07)	0.098
ALT (per 10 unit increase)	0.98 (0.93–1.02)	0.224	0.98 (0.94–1.02)	0.302
GGT (per 10 unit increase)	0.99 (0.88–1.16)	0.854	1.08 (0.89–1.19)	0.879
Proteinuria	1.07 (0.75–1.41)	0.631	1.06 (0.71–1.37)	0.878
Oral hygiene care				
Frequency of tooth brushings (times/day)				
0–1	1 (Reference)		1 (Reference)	
2	0.81 (0.68–0.94)	0.003	0.94 (0.79–1.12)	0.520
≥3	0.48 (0.41–0.59)	<.001	0.74 (0.60–0.90)	0.002
Dental visits for any reason	0.86 (0.73–0.99)	0.046	1.01 (0.87–1.18)	0.864
Professional teeth scaling	0.64 (0.52–0.78)	<.001	0.94 (0.77–1.15)	0.527
Oral health status				
Periodontal disease	0.76 (0.51–1.02)	0.058	0.73 (0.58–1.02)	0.068
Number of teeth lost				
0	1 (Reference)		1 (Reference)	
1–7	1.37 (1.21–1.69)	<.001	1.38 (1.19–1.61)	<.001
8–14	2.42 (2.21–3.97)	<.001	1.60 (1.14–2.24)	0.005
≥15	3.61 (2.67–5.18)	<.001	1.81 (1.31–2.51)	<.001

Multivariable model was adjusted for age, sex, income level, alcohol consumption, smoking

status, regular exercise, body mass index (kg/m^2), hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria, periodontal disease, frequency of tooth brushings, dental visits for any reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

HR, hazard ratio; CI, confidence interval; SBP, systolic blood pressure; FBS, fasting blood sugar; AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, gamma glutamyl transferase.

Table S10. Subgroup analysis regarding the occurrence of oral cavity and oral cancer.

Variables	Number of events	ICD-10 code (C00-C06); oral cavity and oral cancer			
		Follow-up duration (person-years)	Age-adjusted incidence rate (per 1000 person-yrs, (95% CI))	Multivariable analysis	
				HR (95% CI)	p value
Periodontal disease					
No	812	1500392	0.38 (0.35–0.42)	1 (Reference)	
Yes	82	225419	0.26 (0.20–0.32)	0.73 (0.58–1.02)	0.068
Frequency of tooth brushings (times/day)					
0–1	189	248527	0.43 (0.37–0.50)	1 (Reference)	
2	471	774947	0.41 (0.37–0.46)	0.94 (0.79–1.12)	0.520
≥3	234	702337	0.29 (0.26–0.34)	0.74 (0.60–0.90)	0.002
P for trend				0.001	
Dental visits for any reason					
No	560	996628	0.39 (0.35–0.43)	1 (Reference)	
Yes	334	729183	0.34 (0.30–0.38)	1.01 (0.87–1.18)	0.864
Professional teeth scaling					
No	752	1312773	0.39 (0.35–0.42)	1 (Reference)	
Yes	142	413038	0.31 (0.26–0.36)	0.94 (0.77–1.15)	0.527
Number of teeth lost					
0	562	1315230	0.33 (0.30–0.37)	1 (Reference)	
1–7	250	370512	0.46 (0.40–0.53)	1.38 (1.19–1.61)	<.001
8–14	38	24313	0.58 (0.42–0.81)	1.60 (1.14–2.24)	0.005
≥15	44	15756	0.69 (0.50–0.95)	1.81 (1.31–2.51)	<.001
P for trend				<.001	

Multivariable model was adjusted for age, sex, income level, alcohol consumption, smoking status, regular exercise, body mass index, hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria, periodontal disease, frequency of tooth brushings, dental visits for any

reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

ICD-10, International Statistical Classification of Diseases, 10th Revision; HR, hazard ratio; CI, confidence interval.

*: The trend test for hazard ratios.

Multivariable analysis was performed by using regression methods of Fine and Gray for competing risk data (death is a competing event for oral cavity and oral cancers).

Table S11. Subgroup analysis regarding the occurrence of salivary gland cancer.

Variables	Number of events	ICD-10 code (C07-C08); salivary gland cancer			
		Follow-up duration (person-years)	Age-adjusted incidence rate (per 1000 person-yrs, (95% CI))	Multivariable analysis	
				HR (95% CI)	p value
Periodontal disease					
No	66	1506405	0.04 (0.03–0.05)	1 (Reference)	
Yes	9	226019	0.04 (0.02–0.07)	0.81 (0.40–1.64)	0.552
Frequency of tooth brushings (times/day)					
0–1	14	249965	0.05 (0.03–0.08)	1 (Reference)	
2	32	778467	0.04 (0.03–0.06)	0.82 (0.44–1.52)	0.523
≥3	29	703992	0.04 (0.03–0.06)	0.93 (0.49–1.78)	0.828
P for trend				0.962	
Dental visits for any reason					
No	37	1000793	0.04 (0.03–0.05)	1 (Reference)	
Yes	38	731631	0.05 (0.04–0.07)	1.76 (1.09–2.85)	0.020
Professional teeth scaling					
No	62	1318351	0.04 (0.03–0.06)	1 (Reference)	
Yes	13	414073	0.03 (0.02–0.06)	0.55 (0.29–1.02)	0.057
Number of teeth lost					
0	48	1319413	0.04 (0.03–0.05)	1 (Reference)	
1–7	24	372346	0.06 (0.04–0.09)	1.62 (0.99–2.67)	0.056
8–14	0	24603	0	Non-estimable	
≥15	3	16061	0.12 (0.04–0.40)	3.64 (1.05–12.56)	0.041

Multivariable model was adjusted for age, sex, income level, alcohol consumption, smoking status, regular exercise, body mass index, hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria, periodontal disease, frequency of tooth brushings, dental visits for any reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

ICD-10, International Statistical Classification of Diseases, 10th Revision; HR, hazard ratio; CI, confidence interval.

*: The trend test for hazard ratios.

Multivariable analysis was performed by using regression methods of Fine and Gray for competing risk data (death is a competing event for salivary gland cancers).

Table S12. Subgroup analysis regarding the occurrence of tonsil and pharynx cancer.

Variables	Number of events	ICD-10 code (C09-C14); tonsil and pharynx cancer			
		Follow-up duration (person-years)	Age-adjusted incidence rate (per 1000 person-yrs, (95% CI))	Multivariable analysis	
				HR (95% CI)	p value
Periodontal disease					
No	206	1505683	0.12 (0.10–0.14)	1 (Reference)	
Yes	35	225879	0.13 (0.09–0.18)	1.07 (0.73–1.57)	0.726
Frequency of tooth brushings (times/day)					
0–1	52	249759	0.15 (0.11–0.20)	1 (Reference)	
2	107	778114	0.11 (0.09–0.14)	0.83 (0.59–1.15)	0.263
≥3	82	703689	0.11 (0.09–0.14)	0.87 (0.61–1.25)	0.454
P for trend				0.542	
Dental visits for any reason					
No	138	1000278	0.12 (0.10–0.14)	1 (Reference)	
Yes	103	731285	0.12 (0.10–0.15)	1.11 (0.82–1.50)	0.490
Professional teeth scaling					
No	192	1317691	0.12 (0.10–0.14)	1 (Reference)	
Yes	49	413871	0.12 (0.09–0.15)	0.92 (0.64–1.33)	0.662
Number of teeth lost					
0	175	1318743	0.12 (0.10–0.14)	1 (Reference)	
1–7	57	372177	0.12 (0.09–0.16)	0.94 (0.69–1.27)	0.675
8–14	7	24578	0.14 (0.07–0.31)	1.08 (0.49–2.35)	0.851
≥15	2	16064	0.05 (0.01–0.19)	0.33 (0.08–1.35)	0.124
P for trend				0.233	

Multivariable model was adjusted for age, sex, income level, alcohol consumption, smoking status, regular exercise, body mass index, hypertension, diabetes mellitus, dyslipidemia, systolic blood pressure, total cholesterol, fasting blood sugar, aspartate aminotransferase, alanine aminotransferase, gamma glutamyl transferase, proteinuria, periodontal disease, frequency of tooth brushings, dental visits for any

reason, professional teeth scaling, and number of teeth lost except regarding the independent variable.

ICD-10, International Statistical Classification of Diseases, 10th Revision; HR, hazard ratio; CI, confidence interval.

*: The trend test for hazard ratios.

Multivariable analysis was performed by using regression methods of Fine and Gray for competing risk data (death is a competing event for tonsil and pharynx cancers).