Supplementary table 1: Logistic univariable regression analysis the Association between covariates and early-onset OA using NHANES data (N = 26145,1999-2020)

Variable	OR_95CI	*P*-value
Age (Years)	1.10 (1.09~1.11)	<0.00
Gender: Female vs male	1.52 (1.34~1.72)	<0.00
Race		
Non-Hispanic White	1(reference)	
Non-Hispanic Black	0.56 (0.48~0.66)	<0.00
Mexican American	0.27 (0.21~0.34)	<0.00
Other	0.56 (0.47~0.66)	<0.00
Educational level (Years)		
<9	1(reference)	
9-12	2.06 (1.43~2.95)	<0.00
>12	2.45 (1.72~3.49)	<0.00
Marital Status		
Married or living with partners	1(reference)	
Living alone	0.99 (0.87~1.12)	0.82
PIR		
Low income	1(reference)	
Medium income	0.89 (0.76~1.03)	0.12
High income	1.04 (0.89~1.21)	0.62
TC (mmol/L)	1.13 (1.07~1.19)	<0.00
HDL-C (mmol/L)	0.94 (0.81~1.09)	0.43
Serum calcium(mg/dl)	0.83 (0.71~0.98)	0.03
Serum phosphorus(mg/dl)	1.01 (0.91~1.13)	0.80
BUN (mg/dl)	1.04 (1.02~1.05)	<0.00
UA (mg/dl)	1.05 (1~1.09)	0.04
BMI (kg/m2)	1.05 (1.05~1.06)	<0.00
Alcohol use: Yes vs No	1.38 (1.17~1.62)	<0.00
Hypertension: Yes vs No	3.08 (2.71~3.49)	<0.00
Diabetes: Yes vs No	2.70(2.24~3.26)	<0.00
CHD: Yes vs No	3.68 (2.42~5.59)	<0.00
Stroke: Yes vs No	4.04 (2.88~5.66)	<0.00

Abbreviations: OR, odds ratio.CI,Confidence Interva.PIR, povertyto-income ratio. BMI, body mass index.TC,total cholesterol.HDL-C,high-density lipoprotein cholesterol.BUN, blood urea nitrogen.UA, uric acid.CHD,coronary heart disease.

Supplementary table 2: Subgroup analysis of the association between tobacco smoke exposure and early-onset OA using NHANES data (N=26145,1999-2020)

Subgroup	N	Event (%)	aOR (95%CI)	*P* for interaction
Age				
<40years	15315	258 (1.70)	1.07 (1.04~1.11)	0.13
≥40years	10830	828 (7.60)	1.05 (1.03~1.07)	
Gender				
Male	12596	417 (3.30)	1.05 (1.02~1.08)	0.55
Female	13549	669 (4.90)	1.06 (1.03~1.08)	
Marital status				
Married or living with a partner	16069	671 (4.20)	1.05 (1.03~1.08)	0.75
Living alone	10076	415 (4.10)	1.06 (1.03~1.08)	
BMI				
<25kg/m^2	8212	212 (2.60)	1.04 (1~1.08)	0.06
25-30kg/m^2	8282	290 (3.50)	1.08 (1.04~1.11)	
≥30kg/m^2	9651	584 (6.10)	1.04 (1.02~1.07)	
Alcohol use				
No	5739	186 (3.20)	1.09 (1.04~1.14)	0.16
Yes	20406	900 (4.40)	1.05 (1.03~1.07)	
Hypertension				
No	21147	645 (3.10)	1.05 (1.03~1.07)	0.52
Yes	4998	441 (8.80)	1.06 (1.03~1.09)	
Diabetes				
No	24726	948 (3.80)	1.06 (1.04~1.08)	0.13
Yes	1419	138 (9.70)	1.01 (0.96~1.06)	

Abbreviations: aOR,adjusted odds ratio.CI, confidence interval.BMI, body mass index.

Except for the stratification component itself, each stratification factor was adjusted for all other variables (Age+Gender+Race+Educational level+marital+PIR+TC+HDL-C+Serum calcium+Serum phosphorus+BUN+UA+BMI+Alcohol use+Hypertension+Diabetes+CHD+Stroke).

An observational study is a type of epidemiological study design, which can take the form of a cohort, a case–control, or a cross-sectional study. When presenting observational studies in manuscripts, an author needs to ascertain a clear presentation of the work and provide the reader with appropriate information to enable critical appraisal of the research. The Strengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines were created to aid the author in ensuring high-quality presentation of the conducted observational study. The original articles publishing the STROBE guidelines together with their bibliographies were identified and thoroughly reviewed. These guidelines consist of 22 checklist items that the author needs to fulfil before submitting the manuscript to a journal. The STROBE guidelines were created to aid the authors in presenting their work and not to act as a validation tool for the conducted study or as a framework to conduct an observational study on. The authors complying with these guidelines are more likely to succeed in publishing their observational study work in a journal.

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	**Recommendation**
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
Outcome data	15*	Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included

		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.