Supplementary file 1

STROBE statement — Checklist of items for cross-sectional studies

Tobacco-related knowledge, attitudes and practices among college students in Fujian Province, China: A cross-sectional study

Section/Topic	Item No.	Recommendation	Page/Section	Reported on
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	Title, p.1; Abstract	Title includes "cross-sectional study"; structured Abstract with INTRODUCTION, METHODS, RESULTS, CONCLUSIONS (293 words, within 300-word limit)
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction, ¶1–5	Tobacco burden in China, Fujian Province data gaps, e-cigarette trends, KAP framework rationale, policy context
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction, final ¶	Three numbered study aims: (1) KAP assessment, (2) associated factors, (3) channel–knowledge association
Methods				
Study design	4	Present key elements of study design early in the paper	Methods, opening sentence	STROBE Statement explicitly cited in opening sentence; cross-sectional design stated; full STROBE checklist provided as Supplementary File 1
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Methods, ¶1	Single university in Fuqing City, Fujian Province; data collected December 2025; Wenjuanxing online survey platform
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants	Methods, ¶1	Eligibility: undergraduate students aged ≥18 years enrolled at the institution; convenience plus snowball sampling distributed via class

Section/Topic	Item No.	Recommendation	Page/Section	Reported on
				groups and the university's social-media channels; exclusion criteria (completion <3 min, non-undergraduate, >50% missing on tobacco items)
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers	Methods, Measures	KAP dimensions defined with explicit response options for every item: Knowledge (4 disease items, SHS, e-cigarette awareness, addictiveness, attractiveness, with scoring rule for the 0–4 knowledge score), Attitude (10 venue bans, tourism ban, tax, low-tar perception, home rules), Practice (ever use, quit attempts, susceptibility), 8 information channels, social environment, demographics with categories specified
Data sources / measurement	8*	For each variable of interest, give sources of data and details of methods of assessment	Methods, Measures	Instrument developed with reference to the China CDC 2021 National College Student Tobacco Survey, the Global Adult Tobacco Survey (GATS) Core Questionnaire (citation 29), and the Global Youth Tobacco Survey (GYTS) (citation 30); response options stated for every item; 5-point Likert scale dichotomized as agree (4–5) vs. other (1–3); knowledge items characterized as independent factual indicators; pilot tested with 30 students (excluded from analysis); content validity reviewed by 2 public-health faculty
Bias	9	Describe any efforts to address potential sources of bias	Methods, ¶1; Discussion, Limitations	Exclusion of rapid completers (<3 min); anonymous voluntary participation; sampling bias (convenience and snowball, single-institution origin, female overrepresentation, online-only delivery) acknowledged

Section/Topic	Item No.	Recommendation	Page/Section	Reported on
				as the most important limitation; social desirability bias and other limitations discussed
Study size	10	Explain how the study size was arrived at	Methods, ¶1–2	A priori sample-size estimation using $n = Z^2 \times p \times (1-p) / d^2$ with Z=1.96, expected prevalence p=12% (based on the 2021 China CDC national college survey), and margin of error d=4%, yielding minimum n=254; recruitment target 360 (allowing 30% invalid responses); achieved 492 returned, 77 excluded, 415 valid (response rate 84.35%)
Quantitative variables	11	Explain how quantitative variables were handled in the analyses	Methods, Statistical analysis	Knowledge score 0–4 (sum of correct "yes" responses across four disease-specific items, with "no" and "don't know" each scored 0); channel count 0–8 categorized as low (0–5) / medium (6–7) / high (8); Likert items dichotomized as agree (4–5) vs. other (1–3)
Statistical methods	12	(a) Describe all statistical methods (b) Describe methods for subgroups/interactions (c) Explain how missing data were addressed (d) Describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	Methods, Statistical analysis	(a) Chi-squared, Welch t-tests, one-way ANOVA with Bonferroni correction; univariate (crude) and multivariable logistic regression with cORs and aORs reported with 95% CIs; analyses performed in Python 3.11 with SciPy 1.12 / NumPy 1.26 (citation 31) (b) Subgroup analyses by sex, tobacco-use status (c) Exclusion of >50% missing on tobacco items at the screening stage (d) Convenience and snowball sampling acknowledged; 5 multivariable covariates selected a priori and entered simultaneously (e) Not applicable; events-per-variable=9.4 explicitly noted as a limitation

Section/Topic	Item No.	Recommendation	Page/Section	Reported on
Results				
Participants	13*	(a) Report numbers at each stage (b) Give reasons for non-participation (c) Consider use of a flow diagram	Results, ¶1; Methods	(a) 492 returned → 77 excluded → 415 analyzed (b) 41 rapid completers, 25 non-undergraduate, 11 missing data (c) Numbers reported in text; no flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (b) Indicate number with missing data	Results, Participant characteristics; Table 1	(a) Sex, grade, residence, expenditure, social environment reported in Table 1 (b) Missing data handled by exclusion criteria; no additional missing reported in analyzed sample
Outcome data	15*	Report numbers of outcome events or summary measures	Results, Tobacco-use prevalence	47/415 (11.3%) ever tobacco use; 66% recent or continuing users; product type breakdown reported; ever use justified as primary outcome to maximise statistical power
Main results	16	(a) Give unadjusted and adjusted estimates with precision (b) Report category boundaries (c) Consider translating relative to absolute risk	Results; Tables 2–5 plus Supplementary Table S1	(a) Crude (cOR) and adjusted (aOR) odds ratios with 95% CIs and exact p-values reported in Table 2 with grouped Univariate / Multivariable column headings; full footnote lists 5 covariates entered simultaneously; McFadden pseudo-R^2=0.190; events=47; events-per-variable=9.4 (b) Channel exposure categorized: 0–5 / 6–7 / 8 (c) Prevalence rates reported as percentages
Other analyses	17	Report other analyses done — e.g., analyses of subgroups and interactions, and sensitivity analyses	Results; Supplementary File 2	Subgroup analyses: knowledge by sex; entertainment-venue ban support by sex and tobacco-use status; tourism ban by sex and tobacco-use status; addictiveness recognition by tobacco-use status; perceived attractiveness of smoking by sex. Per-channel exposure-knowledge t-tests for each of 8 information channels reported

Section/Topic	Item No.	Recommendation	Page/Section	Reported on
				in Supplementary File 2 (Supplementary Table S1)
Discussion				
Key results	18	Summarise key results with reference to study objectives	Discussion, opening ¶	Opening summary paragraph lists four principal findings: cardiovascular vs. lung-cancer knowledge gap; discrepancy between general harm awareness and product-specific risk literacy; lower addictiveness recognition among ever users (consistent with optimistic bias); broad support for smoke-free tourism. Subsections then expand on each theme
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision	Discussion, Limitations	Seven limitations are presented in descending order of importance, beginning with sampling bias as the most important: (i) substantial sampling bias from convenience/snowball recruitment, single-institution origin, female overrepresentation (75.4%), and online-only delivery; (ii) cross-sectional design and reverse-causality concerns for the channel–knowledge association; (iii) self-reported data and social-desirability bias; (iv) low events-per-variable ratio of 9.4; (v) binary channel measurement; (vi) only 4 disease-specific knowledge items; (vii) no formal psychometric validation. Strengths also stated
Interpretation	20	Give a cautious overall interpretation considering objectives, limitations, and other evidence	Discussion	Findings contextualized with GATS China 2018, Hebei, Ningxia, Hangzhou, and Fuzhou studies; interpretive language consistently moderated to "consistent with", "may inform", "may

Section/Topic	Item No.	Recommendation	Page/Section	Reported on
				merit further evaluation" given the cross-sectional design; reverse causality explicitly noted for channel–knowledge association; hypothetical-scenario limitation explicitly noted for tourism-ban data
Generalisability	21	Discuss the generalisability (external validity) of the study results	Discussion, Limitations	Single university with 75.4% female enrollment; convenience/snowball sampling; generalizability explicitly limited; not representative of Fujian Province or national trends. The reordered Limitations paragraph identifies sampling bias as the dominant generalisability concern
Other information				
Funding	22	Give the source of funding and the role of the funders	Funding statement	Humanities and Social Sciences Youth Foundation of the Ministry of Education of China (No. 25XJCZH009); 2025 Undergraduate Education and Teaching Research Project of Fujian Polytechnic Normal University (XJBJY13). Funders had no role in study design, data collection, analysis, or manuscript preparation

** Give information separately for exposed and unexposed groups.*

Supplementary file 2

Manuscript ID: TID-01943-2026-01

Tobacco-related knowledge, attitudes, and practices among college students in Fujian Province, China: A cross-sectional study

Supplementary table S1. Tobacco-control information-channel exposure (past 30 days) and association with disease-specific health-knowledge score among undergraduate students at a university in Fuqing City, Fujian Province, China: a cross-sectional online survey, December 2025 (N=415).

Channel	Exposure n (%)	Knowledge score, mean ± SD		*t*	p-value
		Exposed	Unexposed		
Social media platforms (e.g., WeChat, Weibo, Xiaohongshu)	345 (83.1)	3.08 ± 1.27	2.10 ± 1.63	4.14	p<0.001
School health education (e.g., campus medical centre, lectures)	342 (82.4)	3.06 ± 1.28	2.38 ± 1.63	2.78	p=0.007
Billboards / posters / printed materials	338 (81.4)	3.09 ± 1.27	2.15 ± 1.62	4.08	p<0.001
Short-video platforms (e.g., Douyin, Kuaishou, Shipinhao)	335 (80.7)	3.08 ± 1.28	2.28 ± 1.60	3.45	p=0.001
Medical institution platforms (e.g., hospital websites, official accounts)	321 (77.3)	3.12 ± 1.25	2.29 ± 1.61	3.86	p<0.001
Traditional media (television, radio, newspapers, magazines)	320 (77.1)	3.14 ± 1.24	2.33 ± 1.59	3.95	p<0.001
Family / friends	308 (74.2)	3.12 ± 1.26	2.37 ± 1.56	3.81	p<0.001
Professional health websites / apps (e.g., DXY, 39 Health, Chunyu)	293 (70.6)	3.19 ± 1.22	2.33 ± 1.56	4.55	p<0.001

Disease-specific health-knowledge score: sum of correct ("yes") responses across four items (smoking causes stroke, heart disease, lung cancer, erectile dysfunction); range 0-4. Exposure was assessed for the past 30 days as a binary variable (yes/no) for each channel. Group means were compared with independent-samples t-tests (Welch's correction; two-tailed). Channels are ordered by exposure rate from highest to lowest. All eight channels showed a statistically significant positive association with the health-knowledge score at p<0.01; the corresponding total channel-count dose-response analysis is presented as Table 5 in the main manuscript.

Notes

The eight channels listed in Supplementary table S1 were not pre-specified for multiple-comparison correction because they constitute distinct exposure indicators rather than tests of a single hypothesis; the corresponding dose-response analysis using the total channel count (Table 5 in the main manuscript) provides the principal inferential test of the channel-exposure-knowledge relationship and uses Bonferroni correction across the three exposure-count strata. Effect sizes (Cohen's d) for each per-channel comparison ranged from approximately 0.43 to 0.62, indicating small-to-medium effects that are consistent across all eight channels.

As discussed in the main manuscript Limitations, channel exposure was measured at the binary (yes/no) level for the past 30 days; frequency, intensity, and duration of exposure were not captured. The cross-sectional design

precludes causal inference: students with higher health-knowledge scores may both seek out more channels and retain more information from each channel, and reverse causality cannot be excluded.