Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Uptake of recommended oral-health self-care and its social, behavioural and contextual correlates among school-going adolescents in African countries: a theory-informed multi-country analysis of Global School-based Student Health Survey data

Wingston Felix Ng'ambi

wingston.ngambi@gmail.com

Kamuzu University of Health Sciences

Cosmas Zyambo

University of Zambia

Research Article

Keywords: oral health, toothbrushing, adolescents, Africa, GSHS, complex survey analysis, predictor selection, social determinants, hygiene, parental monitoring, substance use

Posted Date: October 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11102426/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Twice-daily toothbrushing is a simple, scalable oral-health self-care practice, yet African evidence on its social and behavioural patterning remains fragmented. We examined recommended toothbrushing uptake and its demographic, social, behavioural and contextual correlates among school-going adolescents across participating African GSHS surveys.

Methods

We harmonised 20 public-use GSHS survey files from 12 African countries. A valid twice-daily toothbrushing outcome and complete survey design were available for 52,559 adolescents from 17 surveys in 11 countries. Age, sex, relative grade progression and country were protected a priori, while modifiable candidates were pre-specified across socioeconomic, dietary, hygiene, safety, lifestyle, school, family, psychosocial and substance-use domains. Candidate-specific models used maximal valid subsets. Non-composite candidates with at least 60% outcome-sample coverage and representation in at least six countries were eligible for constrained AIC selection on a fixed complete-case common basket. Survey-weighted inference accounted for weights, strata and primary sampling units.

Results

Overall weighted recommended toothbrushing uptake was 66.4% (95% CI 65.4% to 67.4%), with 69.6% among females and 63.6% among males. Male sex was associated with lower uptake (OR 0.75, 95% CI 0.71 to 0.79), as was severe food insecurity (aOR 0.83, 95% CI 0.77 to 0.90). Strong positive clustering was observed for handwashing before eating (aOR 2.85), handwashing after toilet (2.84), soap use (2.75), parental homework checking (1.93) and parental monitoring (1.45). Truancy, parental tobacco use, alcohol use, cigarette smoking and cannabis use were associated with lower uptake. The common-basket model included 7,740 adolescents, evaluated 23 eligible candidates and retained 15; weighted AUC increased from 0.557 to 0.714 and Brier score improved from 0.226 to 0.199.

Conclusions

Recommended toothbrushing uptake was incomplete and patterned across multiple social and behavioural domains. General hygiene self-care and supportive family and school environments clustered with higher uptake, whereas social disadvantage and several risk behaviours clustered with lower uptake. These findings support integrated school oral-health promotion; the high-dimensional common-basket model is exploratory and should not be interpreted causally or as representative of Africa as a whole.

Background

Oral health is integral to general health, nutrition, well-being, learning and social participation. Oral diseases remain among the most prevalent health conditions globally and their burden is unequally distributed, with disadvantage accumulating along social, economic and environmental gradients [1–6]. Dental caries and periodontal disease share determinants with other noncommunicable conditions, including high free-sugar intake, tobacco use, alcohol use and constrained access to preventive resources [3, 4, 8–11, 14]. This makes adolescence an important period for prevention because everyday self-care practices and wider health behaviours are being consolidated.

The World Health Organization recommends twice-daily toothbrushing with fluoride-containing toothpaste as a core preventive measure for dental caries [1, 7]. Toothbrushing frequency is therefore a useful behavioural indicator, although frequency alone does not establish brushing quality, fluoride exposure or clinical oral-health status. In school-aged populations, oral health is also shaped by interacting influences at individual, family, school and community levels [12]. Health-promoting-school approaches are especially relevant because schools can link oral health with hygiene, nutrition, tobacco and alcohol prevention, family engagement and supportive environments [9, 13]. The historical African GSHS surveys available for this study do not consistently contain fluoride-toothpaste use, dentist attendance, toothache or oral-health-related school absence; accordingly, the present outcome is precisely defined as uptake of recommended toothbrushing frequency, not clinical oral-health status or dental-service utilisation.

Previous GSHS analyses have documented substantial cross-national variation in toothbrushing and a recurrent sex gap, with boys more likely than girls to report infrequent cleaning in many settings [15–17]. A Zambian GSHS analysis also linked poor oral hygiene with hunger and substance use, suggesting that toothbrushing is embedded within a wider social and behavioural profile rather than operating as an isolated dental behaviour [18]. However, African evidence remains fragmented across single-country analyses and global studies, and many analyses evaluate only a restricted set of correlates.

We therefore examined a broader set of harmonisable GSHS variables. The objectives were to estimate weighted uptake across demographic and country strata; evaluate socioeconomic, diet, hygiene, safety, activity, school, family, psychosocial and substance-use correlates; fit domain-specific multivariable models; apply constrained predictor selection on a fixed common-basket sample; and assess model stability and interpretability within a complex survey framework.

Methods

Study design and data sources

We conducted a pooled cross-sectional analysis of public-use Global School-based Student Health Survey (GSHS) datasets from participating African countries. GSHS uses school-based cluster sampling and anonymous self-administered questionnaires. Variables were harmonised by question wording and

response meaning rather than variable names or question numbers; historical files did not consistently include the additional oral-health constructs in later GSHS modules [19, 20]. Metadata-aware import, label preservation and provenance checks followed reproducible health-data principles [24]. The geographic coverage of included surveys is shown in Fig. 1.

Study population and sample derivation

The harmonised source contained 61,241 adolescents from 20 survey files in 12 countries. A valid toothbrushing outcome and complete survey-design information were available for 52,559 adolescents from 17 surveys in 11 countries. The protected-core model included 50,912 complete cases after requiring usable age, sex, relative grade and country information. Candidate and domain models used their largest valid subsets; the fixed high-dimensional common basket included 7,740 adolescents from 4 surveys in 3 countries. Sample derivation is summarised in Fig. 2.

Primary outcome

Recommended toothbrushing uptake was defined as cleaning or brushing the teeth at least twice daily during the previous 30 days. No brushing, less than once daily or once daily was coded 0; twice daily or more was coded 1. Missing or non-classifiable responses remained missing. At least once-daily brushing was used as a sensitivity outcome.

Protected demographic and contextual covariates

Age was grouped as 11–13, 14–16 and 17–18 years, with 14–16 years as reference; sex was coded as female or male. Because grade systems differed across surveys, ordered grade was standardised within survey as a relative grade-progression z-score, with original labels retained in the supplementary crosswalk. Country was a protected contextual covariate. Survey period (2000–2004, 2005–2009, 2010–2014, 2015–2019 and 2020–2024) was used descriptively and for sensitivity adjustment where estimable without country-period aliasing.

Candidate predictor domains

Candidate predictors were pre-specified across nine domains: socioeconomic/diet (food insecurity, fruit, vegetables, soft drinks); hygiene (handwashing before eating, handwashing after toilet, soap use); risk/safety (physical attack, fighting, bullying); lifestyle (physical activity, sedentary time, active commuting); school (truancy, kind/helpful peers); family (parental understanding, homework checking, monitoring, tobacco use); psychosocial (loneliness, worry-related sleep loss); and substance use (alcohol, cigarettes, cannabis). Any-current-substance and polysubstance indicators were retained for

secondary analyses but excluded from simultaneous selection because they are deterministic composites.

Conceptual and analytical framework

The analytical framework adapts the predictor-selection principles described by Ng'ambi and colleagues [22] and situates them within social-ecological oral-health, common-risk-factor and health-promoting-school perspectives [9, 12, 13]. Candidate inclusion required theoretical coherence, empirical support, modifiability and cross-cultural validity. Age, sex, relative grade progression and country were protected from automated removal. The framework is analytical rather than causal; associations indicate co-patterning and potential intervention entry points. Figure 3 summarises the framework.

Predictor eligibility and constrained selection

For common-basket selection, non-composite candidates required usable data for at least 60% of outcome-eligible adolescents and representation in at least six countries [22]. A fixed complete-case basket ensured that all AIC comparisons used the same respondents. Age, sex, relative grade progression and country were retained throughout constrained bidirectional selection; candidate-specific and domain models remained primary for broader-coverage inference.

Survey design and statistical analysis

Original sampling weights, strata and primary sampling units (PSUs) were retained, with survey-specific identifiers appended to strata and PSUs so clusters remained unique after pooling. Weights were normalised within survey by their survey-specific mean to harmonise raw weight scales while preserving within-survey relative weights. Weighted prevalence and 95% confidence intervals used Taylor-linearised variance estimation [25, 26].

Survey-weighted logistic regression estimated odds ratios for recommended toothbrushing uptake, following transparent specification and reporting principles [23]. The protected core included age, sex, relative grade progression and country. Candidate models added one pre-specified predictor at a time; domain models fitted variables within each conceptual domain. Constrained AIC selection was performed on the fixed common-basket sample, and the selected model was refitted with design-based variance estimation. Odds ratios are reported with 95% confidence intervals and interpreted as cross-sectional associations, not causal effects.

Model validity and reliability

Model assessment included convergence, events per parameter, matrix conditioning, weighted AUC and weighted Brier score. On the common basket, the protected-core, full-candidate and selected models

were compared using weighted AIC, AUC and Brier score; AIC was used only for within-sample model comparison. Event support was treated as a stability diagnostic rather than a rigid threshold [27, 28].

Missing data and reporting

Complete-case analysis was used within each prespecified model. Module absence was distinguished from respondent-level missingness and was not recoded as absence of exposure. Survey-by-predictor availability and reproducibility documentation are provided in the supplementary package [24]; reporting follows STROBE [21].

Ethics

The analysis used de-identified public-use GSHS datasets; no new participant recruitment, contact or identifiable data collection occurred. Ethical approval and consent/assent procedures applied to the original country surveys under relevant institutional and national requirements. No additional informed consent was sought for this secondary analysis of anonymised public-use data.

Results

Predictor availability and common-basket selection

Among 61,241 source records, 52,559 adolescents contributed to the weighted primary outcome analysis. The harmonisation evaluated 26 candidate variables, with availability varying by module (Fig. 4). Relative grade was available for nearly the full outcome sample, whereas many school, family, hygiene and behavioural measures were available for approximately two-thirds to four-fifths of respondents.

Twenty-three non-composite candidates met the common-basket eligibility rule. Although each candidate individually met the cross-country availability criterion, their joint complete-case overlap was limited to 7,740 adolescents from 4 surveys in 3 countries. Constrained selection retained 15 candidates while keeping age, sex, relative grade progression and country in the model. The selection path is shown in Fig. 5 and model fit is compared in Table 1.

Retained candidates were adequate soap use, handwashing after toilet, daily fruit, parental homework checking, daily soft-drink consumption, handwashing before eating, daily vegetables, kind/helpful school peers, truancy, recent bullying, parental monitoring, current alcohol use, worry-related sleep loss, physical fighting and food insecurity. Statistical retention does not imply independent significance or causality; correlated behaviours may attenuate or change direction when fitted simultaneously.

Table 1
Model fit comparison on the identical common-basket sample

Model	N	Parameters	Weighted AIC	Weighted AUC	Weighted Brier
Protected core	7,740	7	9960.6	0.557	0.226
Full eligible candidate set	7,740	31	9046.8	0.715	0.199
Constrained selected model	7,740	23	9037.2	0.714	0.199

On the identical common-basket sample, weighted AIC declined from 9960.6 for the protected core to 9037.2 for the selected model; weighted AUC increased from 0.557 to 0.714 and Brier score decreased from 0.226 to 0.199 (Table 1). The selected model contained 23 fitted parameters because the protected core was retained and some categorical predictors contributed more than one coefficient.

Model validity, reliability and sensitivity analyses

All principal models converged. The protected-core model included 50,912 adolescents (weighted AUC 0.583; Brier score 0.218). The constrained selected model included 7,740 observations (weighted AUC 0.714; Brier score 0.199; events per parameter 120.2; condition number 11.9), indicating adequate numerical support without severe matrix instability.

The male disadvantage was also observed across all 11 country-specific models. Random-effects synthesis gave a pooled male-versus-female OR of 0.75 (95% CI 0.67 to 0.84; I^2 = 84.1%) [29, 30], and leave-one-country-out estimates remained below 1.0. The at-least-once-daily sensitivity outcome showed a similar sex pattern, supporting robustness while indicating contextual heterogeneity.

Weighted prevalence and social patterning

Overall weighted uptake of recommended toothbrushing was 66.4% (95% CI 65.4% to 67.4%). Uptake was higher among females than males and modestly higher in the oldest age group. Country estimates ranged from 57.6% in Algeria to 80.6% in Malawi (Fig. 1). Survey-period estimates were 63.5% in 2000–2004, 68.9% in 2005–2009, 61.3% in 2010–2014 and 72.8% in 2015–2019; these contrasts reflect changing country composition as well as time. Table 2 presents the corresponding social and demographic estimates; denominators vary where individual items were missing.

Table 2
Weighted recommended-toothbrushing uptake by selected social and demographic characteristics

Characteristic	Category	N	Weighted uptake, % (95% CI)
Age Group	11–13	11,206	64.8 (63.3–66.3)
Age Group	14–16	32,027	65.7 (64.6–66.9)
Age Group	17–18	8,769	71.2 (69.5–72.8)
Sex	Female	25,631	69.6 (68.5–70.7)
Sex	Male	26,126	63.6 (62.4–64.7)
Food Insecurity	Never/rarely	24,701	67.3 (66.0-68.5)
Food Insecurity	Sometimes	12,650	66.1 (64.7–67.6)
Food Insecurity	Most/always	5,656	64.6 (62.9–66.3)
Survey Period	2000–2004	6,284	63.5 (60.7–66.3)
Survey Period	2005–2009	14,249	68.9 (66.9–70.9)
Survey Period	2010–2014	19,148	61.3 (59.5–63.1)
Survey Period	2015–2019	12,878	72.8 (71.2–74.4)

Protected demographic core

In the protected-core model, male students had 25% lower odds of recommended uptake than female students (OR 0.75, 95% CI 0.71 to 0.79). Adolescents aged 17–18 years had higher odds than those aged 14–16 years (OR 1.16, 95% CI 1.06 to 1.28). Relative grade progression was not independently associated after age, sex and country adjustment, while country effects remained substantial.

Extended candidate predictor associations

The expanded candidate analysis identified associations across several domains. Severe food insecurity was associated with lower uptake (aOR 0.83, 95% CI 0.77 to 0.90). The strongest positive associations were observed for handwashing before eating (aOR 2.85), handwashing after toilet (2.84) and soap use (2.75). Parental understanding, homework checking and monitoring were positively associated, whereas parental tobacco use, truancy, physical fighting and physical attack were inversely associated.

Substance-use patterns were consistently adverse: current alcohol use (aOR 0.83, 95% CI 0.77 to 0.89), cigarette smoking (0.75, 0.68 to 0.83) and cannabis use (0.76, 0.65 to 0.89) were associated with lower uptake.

Daily fruit, vegetable and soft-drink consumption were positively associated with reported twice-daily toothbrushing in candidate models. The soft-drink association should not be interpreted as protective for

oral health and may reflect residual contextual or socioeconomic confounding and behavioural clustering. Candidate-specific estimates are summarised in Fig. 6.

Domain-specific multivariable models

Domain models distinguished patterns that persisted after within-domain adjustment. Severe food insecurity remained inversely associated in the socioeconomic/diet model; handwashing before eating, handwashing after toilet and soap use remained strongly positive in the hygiene model. Truancy and physical fighting were inversely associated, while kind/helpful school peers were positive in the risk/activity/school model. Parental understanding, homework checking and monitoring remained positive and parental tobacco use remained adverse in the family/psychosocial model. Simultaneous adjustment attenuated substance-use coefficients, consistent with overlap among these behaviours. Selected coefficients are shown in Table 3; full domain estimates are provided in Additional file 1.

Table 3
Selected coefficients from domain-specific multivariable models

Domain	Predictor	Model N	Adjusted OR (95% CI)	P value
Socioeconomic and diet	food insecurity	20,783	0.85 (0.75–0.96)	0.008
Socioeconomic and diet	daily soft drink	20,783	1.81 (1.66–1.98)	< 0.001
Hygiene self-care	handwash before eating adequate	31,522	2.05 (1.87–2.24)	< 0.001
Hygiene self-care	handwash after toilet adequate	31,522	1.71 (1.57–1.87)	< 0.001
Hygiene self-care	soap use adequate	31,522	2.21 (2.04–2.39)	< 0.001
Risk, activity and school	physical fight 12m	14,990	0.90 (0.82–0.99)	0.024
Risk, activity and school	truancy any 30d	14,990	0.81 (0.75–0.88)	< 0.001
Risk, activity and school	school kind helpful	14,990	1.56 (1.45–1.68)	< 0.001
Family and psychosocial	parent understanding	13,147	1.18 (1.07–1.30)	0.001
Family and psychosocial	parent homework check	13,147	1.64 (1.46–1.85)	< 0.001
Family and psychosocial	parent monitoring	13,147	1.21 (1.08–1.35)	0.001
Family and psychosocial	parent tobacco use	13,147	0.86 (0.76–0.96)	0.008

Framework-based synthesis of findings

Taken together, the demographic core showed a persistent male disadvantage and strong country patterning; food insecurity was inversely associated with uptake; general hygiene showed the strongest positive clustering; supportive school and family environments were favourable; and substance use clustered with lower uptake. The high-dimensional model reinforced this multidomain interpretation, but its three-country common basket limits generalisability. A visual synthesis is provided in Supplementary Figure S1.

Discussion

Principal findings

This analysis broadens the empirical picture of adolescent oral-health self-care. Approximately two-thirds of adolescents met the twice-daily toothbrushing threshold, with patterning across demographic, socioeconomic, hygiene, family, school and risk-behaviour domains. The most consistent positive correlates were other routine hygiene practices and parental engagement; male sex, severe food insecurity, truancy, parental tobacco exposure and current substance use identified groups with lower uptake. These findings position toothbrushing within a broader health-behaviour system.

The expanded analysis also demonstrates why theory-informed variable selection is preferable to mechanical screening. Some behaviours that are not themselves oral-health protective, particularly daily soft-drink consumption, were positively associated with reported toothbrushing. Such results are important because they expose correlated social and reporting structures. Treating them naively as intervention effects would be misleading. A structured framework encourages interpretation within domains, comparison of maximal-coverage and multivariable estimates, and explicit recognition of residual confounding.

Methodological contribution

The methodological upgrade adapts a recently published open-source predictor-selection workflow for complex youth surveys [22] and is reinforced by published guidance on transparent logistic-regression specification [23] and reproducible handling of labelled health-research datasets [24]. Its central contribution is not stepwise selection by itself. Rather, it is the combination of a protected demographic core, theory-based candidate specification, explicit cross-survey availability checks, a fixed common-basket sample for comparable AIC calculations, metadata-aware harmonisation, and design-based inference after selection. This approach balances interpretability and parsimony while reducing the risk that automated selection will remove essential demographic adjustment variables.

The analysis also distinguishes population-wide inference from high-dimensional exploratory modelling. The prevalence and protected-core analyses use broad outcome coverage. Candidate-specific models maximize available data for each exposure. Domain models preserve conceptual structure. The common-basket selected model is intentionally narrower and is used to examine multivariable behavioural clustering rather than to redefine the primary population estimate.

Comparison with previous evidence

The persistent female advantage in toothbrushing uptake is consistent with earlier multi-country GSHS research in low- and middle-income settings [15–17]. The adverse food-insecurity and substance-use patterns also align with previous African adolescent findings linking hunger and risk behaviours with poorer oral-hygiene practices [18]. The present analysis extends this literature by demonstrating co-patterning with general hygiene and parental engagement and by examining school, family, psychosocial and lifestyle domains within a single harmonisation framework.

Public-health implications

The results support an integrated school-health approach grounded in the common-risk-factor and health-promoting-school traditions [9, 13]. Oral-health promotion should not be treated only as messaging about teeth. Programmes can link toothbrushing with hand hygiene, healthy routines, school attendance, parental engagement, tobacco and substance-use prevention, and access to affordable fluoride toothpaste. The observed food-insecurity gradient also reinforces the importance of equity and upstream social determinants in oral-health programming [10, 11]. The large country differences indicate that intervention design should be locally adapted rather than transferred uniformly across settings.

Family and school environments appear particularly relevant. Parental homework checking, understanding and monitoring were consistently associated with higher uptake, while a kind/helpful peer environment was favourable. These variables are not direct dental interventions, but they identify supportive contexts in which preventive routines may be easier to establish and sustain. School-based oral-health programmes could therefore be strengthened by family engagement and broader health-promoting-school strategies.

Strengths and limitations

Strengths include harmonisation of multiple country surveys, preservation of complex survey design, a transparent variable crosswalk, explicit separation of module absence from respondent missingness, a broader theory-informed predictor set, protected demographic adjustment, maximal-coverage candidate and domain models, and a reproducible constrained selection workflow consistent with transparent labelled-data practices [24]. The package includes Python code, an R reference implementation, model outputs, predictor-availability files, selection logs, figures and environment information.

Several limitations are important. First, all measures are self-reported and may be affected by recall or social-desirability bias. Second, cross-sectional associations cannot establish direction or causality. Third, toothbrushing frequency does not capture brushing technique, duration or fluoride toothpaste use. Fourth, survey modules varied, so predictors have different denominators. Fifth, the common-basket selected model is based on only three countries and four surveys, limiting external generalisability. Sixth, within-survey relative grade progression is an adjustment strategy rather than exact educational-system equivalence. Seventh, AIC selection can retain correlated variables that are not independently significant, so the selected model complements rather than replaces theory-based domain interpretation. Finally, within-survey weight normalisation supports pooled survey analysis but does not create a population-weighted estimate for Africa as a whole.

Conclusions

Recommended twice-daily toothbrushing was not universal among participating African school-going adolescents and was embedded within a wider social and behavioural ecology. Male students, adolescents facing food insecurity, those reporting weaker school or family protection, and several risk-behaviour groups had lower uptake, while general hygiene self-care and parental engagement clustered strongly with better toothbrushing practice. A theory-informed predictor-selection framework added

analytical depth without discarding core confounder control or complex survey design. The findings support integrated school and family oral-health promotion while underscoring the need for country-specific implementation and longitudinal evaluation.

Abbreviations

AIC
Akaike information criterion
AUC
area under the receiver operating characteristic curve
CI
confidence interval
GSHS
Global School-based Student Health Survey
OR
odds ratio
aOR
adjusted odds ratio
PSU
primary sampling unit
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology.

Declarations

Ethics approval and consent to participate

This secondary analysis used de-identified public-use GSHS datasets. No new participant recruitment, contact or identifiable data collection occurred. Ethical approval and consent/assent procedures were obtained for the original country surveys under the requirements of the relevant implementing institutions and national authorities. No additional informed consent was sought for this analysis of anonymised public-use data.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Supplementary material

Supplementary Figure S1. Visual summary of geographic coverage, observed toothbrushing trends, expanded candidate associations and constrained predictor selection.

Funding

No specific funding is declared for this study.

Author Contribution

WFN conceived the study, developed the analytical framework, harmonised the data, conducted the statistical analyses, prepared the visualisations, interpreted the findings and drafted the manuscript. CZ contributed to methodological development, interpretation of findings and critical revision of the manuscript. Both authors read and approved the final manuscript

Acknowledgements

The authors acknowledge the World Health Organization, participating countries, survey teams, schools and students whose GSHS data made this secondary analysis possible.

Data Availability

We conducted a pooled cross-sectional analysis of public-use Global School-based Student Health Survey (GSHS) datasets from participating African countries. GSHS uses school-based cluster sampling and anonymous self-administered questionnaires. Variables were harmonised by question wording and response meaning rather than variable names or question numbers; historical files did not consistently include the additional oral-health constructs in later GSHS modules [19,20]. Metadata-aware import, label preservation and provenance checks followed reproducible health-data principles [24]. The geographic coverage of included surveys is shown in Figure 1.

References

1. World Health Organization. Oral health. Geneva: WHO; 2025.
2. World Health Organization. Global oral health status report: towards universal health coverage for oral health by 2030. Geneva: WHO; 2022.
3. Peres MA, Macpherson LMD, Weyant RJ, Daly B, Venturelli R, Mathur MR, et al. Oral diseases: a global public health challenge. Lancet. 2019;394(10194):249–60. 10.1016/S0140-6736(19)31146-8.

4. Watt RG, Daly B, Allison P, Macpherson LMD, Venturelli R, Listl S, et al. Ending the neglect of global oral health: time for radical action. Lancet. 2019;394(10194):261–72. 10.1016/S0140-6736(19)31133-X.
5. GBD 2017 Oral Disorders Collaborators. Global, regional, and national levels and trends in burden of oral conditions from 1990 to 2017: a systematic analysis for the Global Burden of Disease 2017 Study. J Dent Res. 2020;99(4):362–73. 10.1177/0022034520908533.
6. Kassebaum NJ, Bernabe E, Dahiya M, Bhandari B, Murray CJL, Marcenes W. Global burden of untreated caries: a systematic review and metaregression. J Dent Res. 2015;94(5):650–8. 10.1177/0022034515573272.
7. World Health Organization. Fluoride toothpaste to prevent dental caries: briefing note. Geneva: WHO; 2022.
8. Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. Bull World Health Organ. 2005;83(9):661–9.
9. Sheiham A, Watt RG. The common risk factor approach: a rational basis for promoting oral health. Community Dent Oral Epidemiol. 2000;28(6):399–406. 10.1034/j.1600-0528.2000.028006399.x.
10. Petersen PE, Kwan S. Equity, social determinants and public health programmes - the case of oral health. Community Dent Oral Epidemiol. 2011;39(6):481–7. 10.1111/j.1600-0528.2011.00623.x.
11. Watt RG. From victim blaming to upstream action: tackling the social determinants of oral health inequalities. Community Dent Oral Epidemiol. 2007;35(1):1–11. 10.1111/j.1600-0528.2007.00348.x.
12. Fisher-Owens SA, Gansky SA, Platt LJ, Weintraub JA, Soobader MJ, Bramlett MD, et al. Influences on children's oral health: a conceptual model. Pediatrics. 2007;120(3):e510–20. 10.1542/peds.2006-3084.
13. Kwan SYL, Petersen PE, Pine CM, Borutta A. Health-promoting schools: an opportunity for oral health promotion. Bull World Health Organ. 2005;83(9):677–85.
14. Moynihan PJ, Kelly SAM. Effect on caries of restricting sugars intake: systematic review to inform WHO guidelines. J Dent Res. 2014;93(1):8–18. 10.1177/0022034513508954.
15. McKittrick TR, Jacobsen KH. Oral hygiene practices among middle-school students in 44 low- and middle-income countries. Int Dent J. 2014;64(3):164–70. 10.1111/idj.12094.
16. Han L, Gao X, Liao M, Yu X, Zhang R, Liu S, et al. Hygiene practices among young adolescents aged 12–15 years in low- and middle-income countries: a population-based study. J Glob Health. 2020;10(2):020436. 10.7189/jogh.10.020436.
17. Das Gupta R, Kothadia RJ, Haider SS, Mazumder A, Akhter F, Siddika N, et al. Toothbrushing frequency among children and adolescents in 72 countries: findings from the Global School-based Student Health Survey. Dent Med Probl. 2024;61(4):495–506.
18. Siziya S, Muula AS, Rudatsikira E. Self-reported poor oral hygiene among in-school adolescents in Zambia. BMC Res Notes. 2011;4:255. 10.1186/1756-0500-4-255.

19. World Health Organization. Global School-based Student Health Survey: Core Questionnaire, 2021 version. Geneva: WHO; 2021.
20. World Health Organization. Global School-based Student Health Survey: Core-Expanded Questions, 2021 version. Geneva: WHO; 2021.
21. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):e296. 10.1371/journal.pmed.0040296.
22. Ng'ambi WF, Zyambo C, Kazembe L. Methodological guidance for predictor variable selection for adolescent smoking outcomes in Global Youth Tobacco Survey using R and Python. Front Appl Math Stat. 2026;12:1895038. 10.3389/fams.2026.1895038.
23. Ng'ambi WF, Zyambo C, Muula AS. A practical guide to key considerations in logistic regression for clinical and public health research: R tutorial. Malawi Med J. 2025;37. Available from: https://www.mmj.mw/?p=13527. Accessed 19 Sep 2026.
24. Ng'ambi WF, Muula AS. A reproducible R workflow to preserve variable and value labels in Stata, SPSS, and SAS datasets for transparent and reproducible health research. Malawi Med J. 2025;37(3):193–206. 10.4314/mmj.v37i3.10.
25. Lumley T. Complex Surveys: A Guide to Analysis Using R. Hoboken: Wiley; 2010.
26. Korn EL, Graubard BI. Analysis of Health Surveys. New York: Wiley; 1999.
27. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;49(12):1373–9. 10.1016/S0895-4356(96)00236-3.
28. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007;165(6):710–8. 10.1093/aje/kwk052.
29. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557–60. 10.1136/bmj.327.7414.557.
30. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177–88. 10.1016/0197-2456(86)90046-2.

Figures

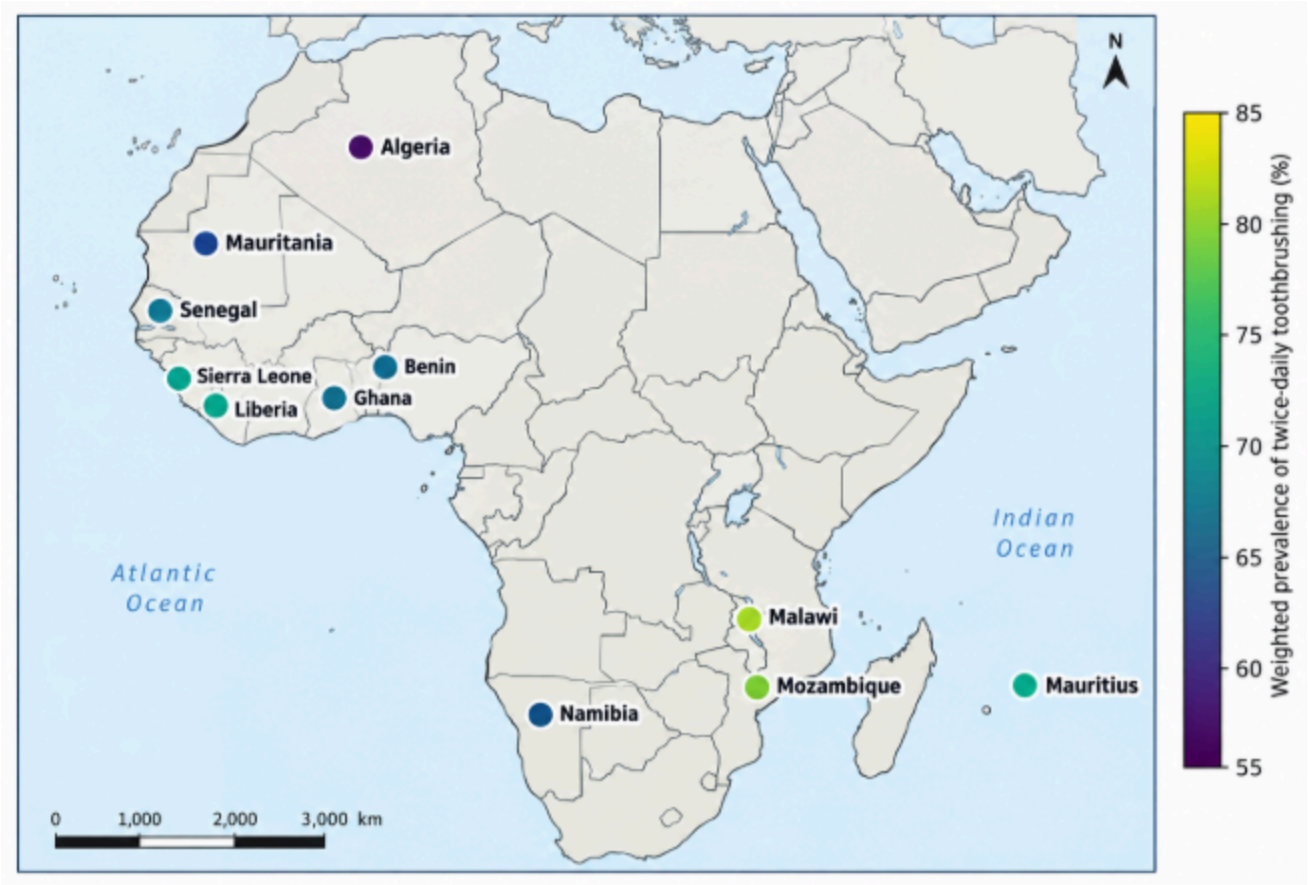


Figure 1

Geographic coverage of the oral-health analysis and country-level weighted prevalence of recommended toothbrushing uptake.

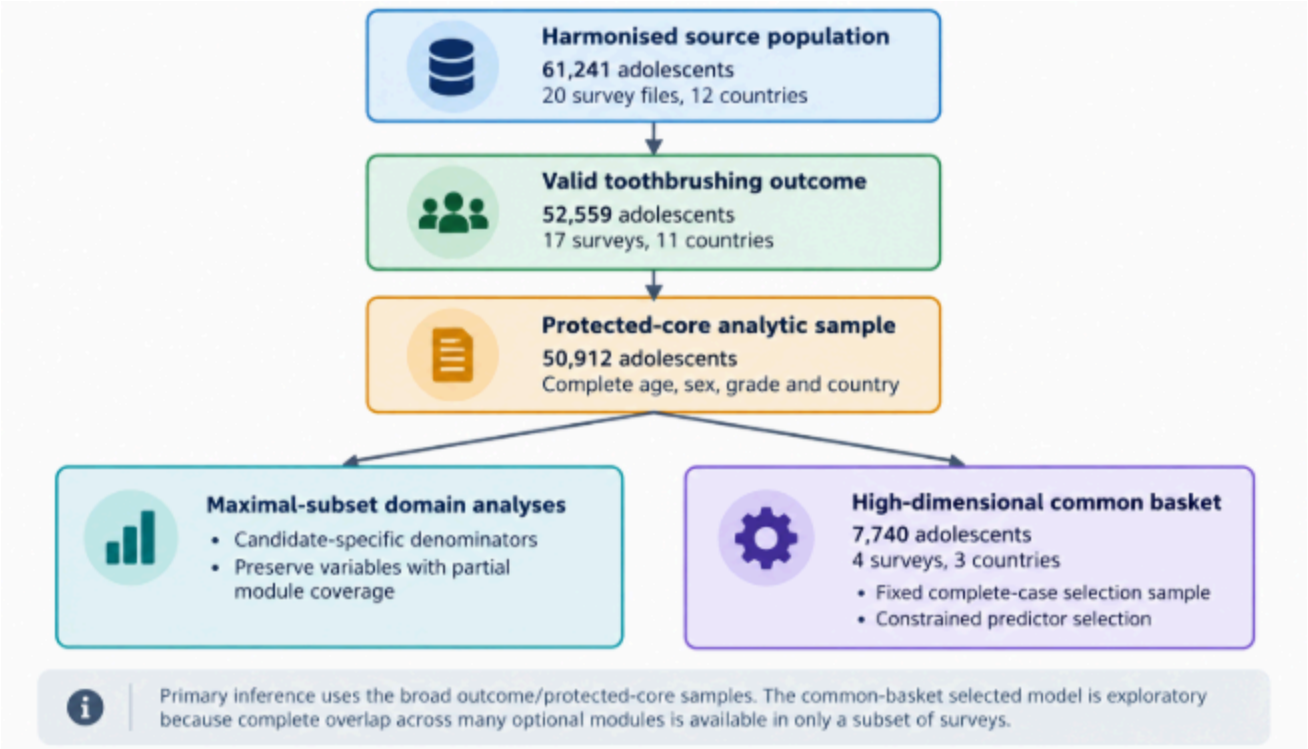


Figure 2

STROBE-style sample flow for the recommended toothbrushing uptake analysis, including the high-dimensional common-basket subset after the population-wide outcome and protected-core samples.

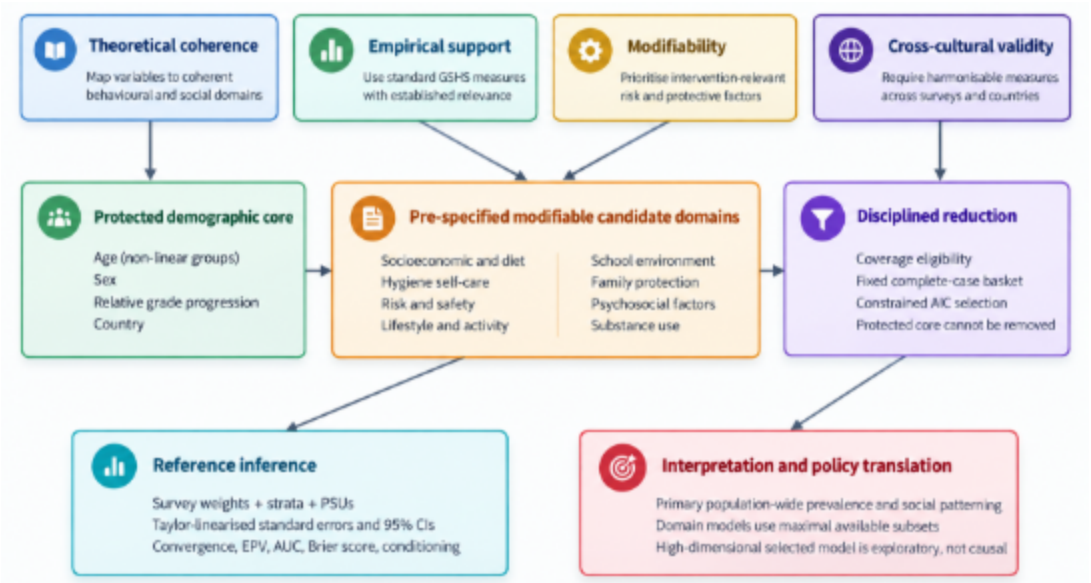


Figure 3

Theory-informed predictor selection and survey-weighted oral-health analysis framework. Candidate variables are pre-specified by domain, screened for cross-survey availability, and reduced using constrained AIC only after defining a fixed complete-case common basket. Final inference retains full survey-design correction.

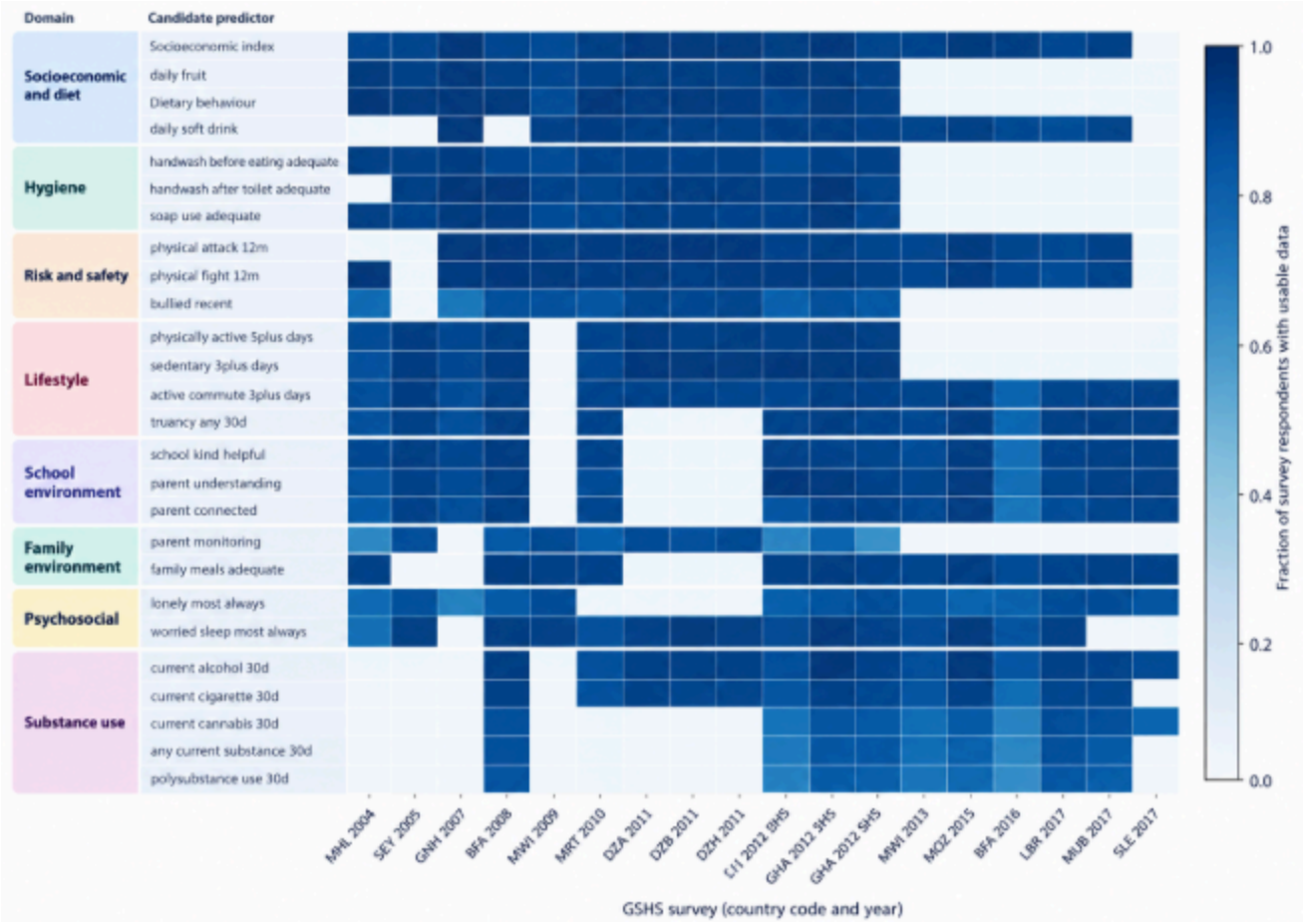


Figure 4

Availability of harmonised candidate predictors across outcome-eligible GSHS surveys. Darker cells indicate higher respondent-level availability within a survey; white cells generally reflect module absence or non-harmonisable content.

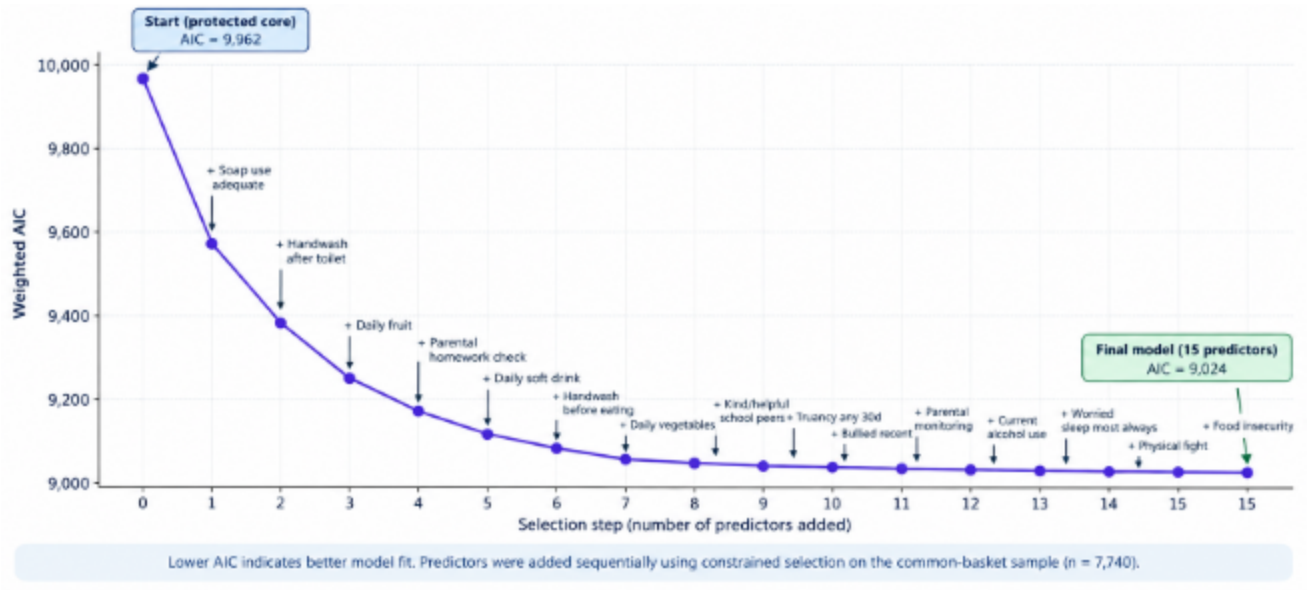


Figure 5

Constrained AIC predictor-selection path on the fixed common-basket sample. The protected demographic core could not be removed.

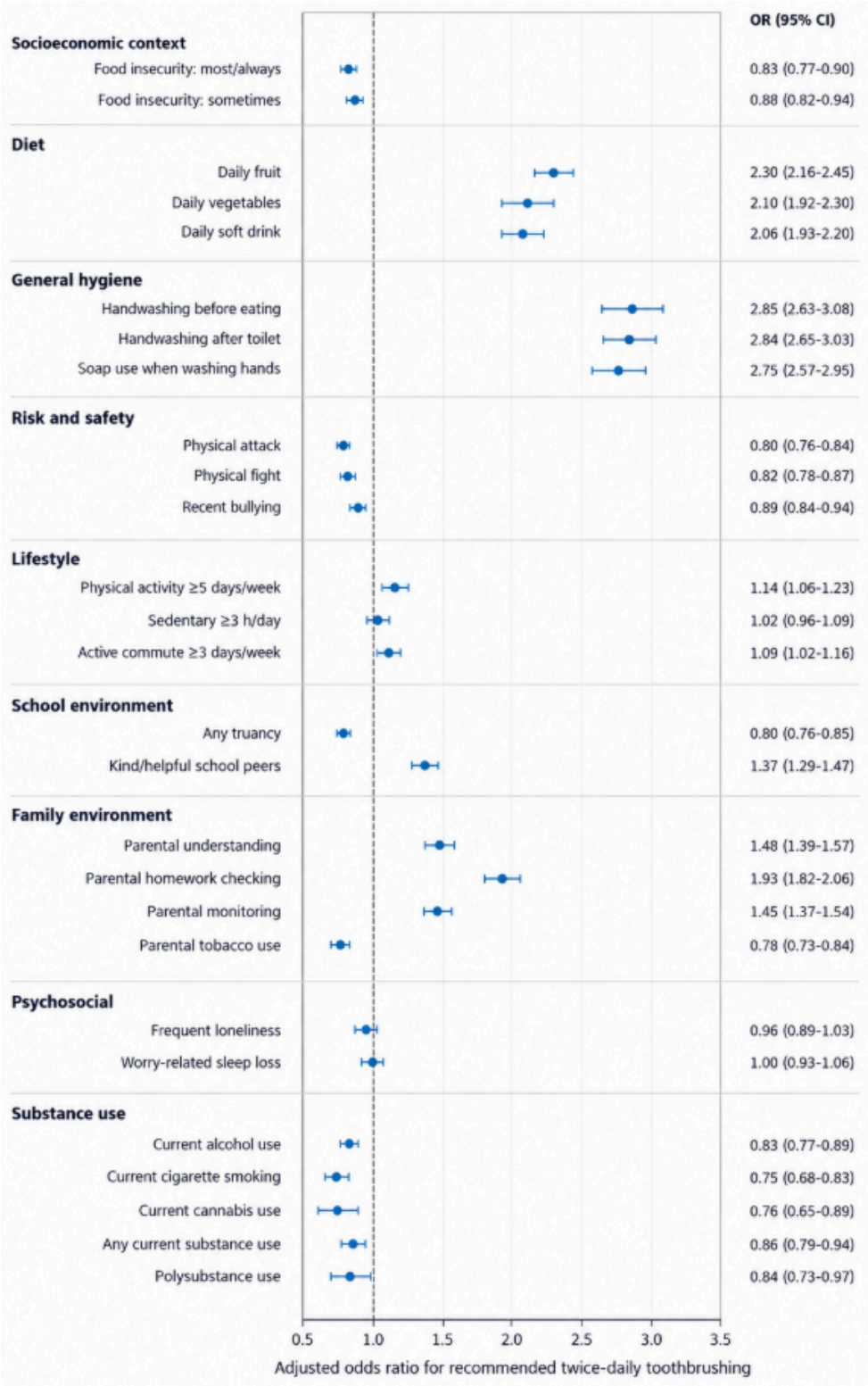


Figure 6

Extended theory-informed candidate predictor associations with recommended twice-daily toothbrushing. Each candidate is fitted in a separate survey-weighted model adjusted for the protected

demographic core and uses its maximal valid subset.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- floatimage7.png