Original Article

Differential impact of socioeconomic position across life on oral cancer risk in Kerala, India: An investigation of life-course models under a time-varying framework

Akhil Soman ThekkePurakkal[1], Ashley Isaac Naimi[2], Sreenath Arekunnath Madathil[1,3], Shahul Hameed Kumamangalam Puthiyannal[1], Gopalakrishnan Netuveli[4], Amanda Sacker[5], Nicolas F Schlecht[6], Belinda Nicolau[1]*

[1]Division of Oral Health and Society, Faculty of Dentistry, McGill University, Montreal, Canada
[2]Department of Epidemiology, University of Pittsburgh, Pennsylvania, United States of America
[3]Epidemiology and Biostatistics Unit, INRS-Institut Armand-Frappier, Laval, Canada
[4]Institute of Health and Human Development, University of East London, London, United Kingdom
[5]UCL Research Department of Epidemiology and Public Health, University College London, United Kingdom
[64]Cancer Prevention and Control, Roswell Park Comprehensive Cancer Center, Buffalo, New York, United States of America

***Corresponding author**
Division of Oral Health and Society - Faculty of Dentistry - McGill University
2001 McGill College, suite 527
Montréal, QC, Canada H3A 1G1
Tel: 514-3987203 ext. 094655 Fax: 514-3987220
Email: belinda.nicolau@mcgill.ca

Abstract

Objectives: The incidence of oral cancer has been rapidly increasing in India, calling for evidence contributing to a deeper understanding of its determinants. Although disadvantageous life-course socioeconomic position (SEP) is independently associated with the risk of these cancers, the explanatory mechanisms remain unclear. Possible pathways may be better understood by testing which life-course model most influences oral cancer risk. We estimated the association between life-course SEP and oral cancer risk under three life-course models: critical period, accumulation and social mobility.

Methods: We recruited incident oral cancer cases (N=350) and controls (N=371) frequency-matched by age and sex from two main referral hospitals in Kozhikode, Kerala, India between 2008 and 2012. We collected information on childhood (0-16 years), early adulthood (17-30 years) and late adulthood (above 30 years) SEP and behavioural factors along the life span using interviews and a life-grid technique. Odds ratios (OR) and 95% confidence intervals (CI) were estimated for the association between life-course SEP and oral cancer risk using inverse probability weighted marginal structural models.

Results: Relative to an advantageous SEP in childhood and early adulthood, a disadvantageous SEP was associated with oral cancer risk [(OR=2.76, 95% CI: 1.99, 3.81) and (OR=1.84, 95% CI: 1.21, 2.79), respectively]. In addition, participants who were in a disadvantageous (vs. advantageous) SEP during all three periods of life had an increased oral cancer risk (OR=4.86, 95% CI: 2.61, 9.06). The childhood to early adulthood social mobility model and overall life-course trajectories indicated strong influence of exposure to disadvantageous SEP in childhood on the risk for oral cancer.

Conclusions: Using novel approaches to existing methods, our study provides empirical evidence that disadvantageous childhood SEP is critical for oral cancer risk in this population from Kerala, India.

1. INTRODUCTION

Oral cancer can be broadly defined as cancers affecting the lip, mouth and parts behind the mouth. It is a disease with low survival rates, and high morbidity, and affects roughly 300,000 people each year, leading to approximately 145,000 deaths worldwide.[1, 2] Developing countries bear two-thirds of the global burden, with India accounting for 25% of new cases and 35% of deaths and where incidence rates have increased considerably in the last decade.[2] A comparison of Globocan data from 2008 and 2012 reveals that the incidence of oral cancer surpassed lung cancer in a span of four years to become the 3rd most common cancer in this country after breast and cervical cancers.[2, 3] Most prevention programmes for these cancers are centred around its strongest risk factors such as paan chewing,[4] bidi and cigarette smoking, and alcohol consumption. [4, 5] However, what has not been emphasised are prevention strategies tailored to socioeconomic contexts critical for comprehensive control of cancers.[6] A major reason for this is the lack of deeper understanding of pathways involving social determinants of health such as socioeconomic position (SEP) and oral cancer.

A cumulative disadvantageous SEP over life has been independently associated with increased risk for this disease.[7] However, SEP varies over the life-course of an individual, a characteristic that is well documented, but consistently overlooked by SEP-oral cancer studies.[8] Appreciating the time-varying nature of SEP provides the unique opportunity to explore the pathways underlying its cumulative effect on oral cancer. The dynamic nature of SEP can be well articulated using the life-course framework, which takes into account the effects of several risk factors spread across multiple points in life.[9] Apart from the cumulative effect of SEP on the outcome (accumulation model), the framework allows to estimate the effect of timing of exposure to disadvantageous SEP at key periods of life (critical period model) that could contribute to the initiation of oral cancer. In addition, the framework permits the investigation of life trajectories generated by the interaction of SEP exposures in multiple periods of life (social mobility model) that can alter an individual's risk of cancer.[9] However, although imperative to its understanding, the relation between SEP and oral cancer has not yet been explored through the lens of multiple life-course models within a single study. Such an investigation may be of special relevance to developing countries such as India, as well as specific populations within India, where high socio-economic disparity exists.

The analysis of life-course models poses challenges. For example, the life-course framework implies that the relation between SEP at several time points and behavioural risk factors are likely subject to complex time-varying feedback loops.[10] Yet, investigators often fail to account for these relations between SEP over the life-course and other time-varying covariates.[7, 11] Therefore, by considering the time-varying nature of SEP and these variables, we estimated the association between SEP measured over three periods of life and oral cancer risk using a case-control study from Kerala, India. We further assessed whether the associations conformed better to a critical period, accumulation or social mobility model.

2. METHODS

Data for this analysis were drawn from the Head and Neck Cancer (HeNCe) Life course study, a multicentre hospital-based case-control study investigating the aetiology of head and neck cancers. Adult participants (N=721) were recruited from the outpatient clinics at two major teaching hospitals, the Government Dental and Medical College and Hospitals, Kozhikode, Kerala, South India between 2008 and 2012. The study design, sample and eligibility criteria have been described in detail elsewhere.[12] Briefly, cases (N=350) included incident, histologically confirmed stage I to IV, consecutive, squamous cell carcinoma cases (C01 and C02: tongue, C03: gum, C04 and C06: floor and unspecified parts of mouth respectively, C05: palate, and C09: tonsil, under International Classification of Diseases 10 Version:2016) of oral cavity diagnosed during the study period. Non-cancer controls (N=371), frequency matched to each identified case by 5-year age group and sex, were randomly selected from 8 outpatient clinics in the same hospitals. Controls were recruited from several clinics (distribution reported elsewhere),[13] not strongly associated with tobacco and alcohol consumption (with no single diagnostic group contributing to more than 20% of the total). This was done to mitigate selection bias.[14] The participation rate was 85.6% and 44.3% among cases and controls respectively.

Data were collected through one-on-one semi-structured interviews using a questionnaire with life-grid technique. Help of a proxy respondent was sought for consenting participants who had difficulty speaking due to disease status. Re-interviews were conducted for 46 randomly selected participants, 6 to 12 weeks after the original interview to test the reliability of the data collected. Ethics approval was obtained from Institutional Review Boards of participating hospitals. Informed written consent was obtained from all participants prior to inclusion in this study.

2.2 Life-course socioeconomic position

Information on housing conditions was used to derive an asset/wealth index that has been documented to be a suitable indicator of SEP for low to middle income societies such as India.[15, 16] We created the wealth index using responses to questions about various assets (housing characteristics, durable assets and access to services),[17] available at the participant's longest place of residence during three time periods: childhood (0-16 years), early adulthood (17-30 years), and late adulthood (above 30 years). Responses to each question were binary coded (Supplemental Appendix file, eTable 1) and a tetrachoric correlation matrix was created for each period (Supplemental Appendix file, eTable 2-4). Principal component analysis was conducted on the correlation matrices and the first component that explained maximum variance (approximately 65%) was extracted.[15] Continuous scores were predicted from these components. The scores for each period were then dichotomized (cut-off at 50th percentile among controls), generating a binary SEP variable (0= advantageous SEP, 1= disadvantageous SEP) for childhood, early adulthood and late adulthood periods each. This variable represented the SEP exposure for each of the three respective critical period models. A four-category variable representing the accumulation model was created by summing the number of periods of disadvantageous SEP (0, 1, 2 and 3). Finally, to test the social mobility models (childhood to early adulthood, and early to late adulthood) we combined the binary SEP variables in respective periods into two variables with four categories representing stable advantageous SEP, upward mobility, downward mobility, and stable disadvantageous SEP. Additional details are provided in Supplemental Appendix file, eAppendix and eTables 1-5.

2.3 Potential confounders

Information on potential confounders and mediators was collected from a set of time-invariant and time-varying factors. The factors included baseline exposures [age (continuous), sex (0: female, 1: male), caste i.e., hierarchy in Hindu religion based on occupation, (0=higher caste, 1=middle caste comprising of backward caste, 2=other backward/scheduled caste/scheduled tribe/others)], education (0=high, 1=low), and time-varying exposures (continuous-cigarette smoking, bidi smoking, paan chewing and alcohol consumption). Education was measured by the number of years of schooling and dichotomized based on the participants' birth cohort (participant's year of birth in our study ranged from 1921 to 1979) to

account for the major social and educational reforms in Kerala in the 1950's.[18] We collected detailed lifetime information on risk behaviours (e.g., duration, quantity, and type of cigarette and bidi smoking, paan chewing, and alcohol consumption) as described elsewhere.[13] This information was used to compute continuous measures of pack-years of cigarette and bidi smoked, chew-years of paan, and number of standard drinks of alcohol per week corresponding to multiple life periods.[13] Additional details are provided in Supplemental Appendix file, eAppendix.

The directed acyclic graph in Figure 1 represents the assumed temporal relations between these variables. Although this is a case-control study, our unique data collection procedure allowed us flexibility to appreciate the temporal relation between vectors representing potential confounders (C0: baseline covariates, C1: 0-16 years, C2a: 17-23 years, C2b: 24-30 years, C3a: 31-50 years, C3b: above 50 years), SEP exposure in the three periods of life and oral cancer. We adjusted for categorical and continuous confounders using indicator coding and restricted cubic splines respectively.

2.4 Statistical methods

T-tests and chi-square tests were used to describe the distribution of continuous and categorical variables, respectively. Our primary aim was to assess the relation between life-course SEP and oral cancer under the three conceptual life-course models. Due to their time-varying nature, SEP and related confounders may also act as mediators. Consequently, standard regression methods may produce biased estimates of exposure-outcome association, regardless of the method used to adjust for confounders. We therefore used inverse probability weighted marginal structural models to account for such confounding and derive our estimates.[19] The inverse probability weighting creates a pseudo reweighted sample where the exposure is independent of the measured potential confounders. We assumed that our case-control data arose from an underlying cohort representing the population of interest.[20]

Weights were derived by fitting a separate exposure model for each period of life and were computed as the inverse of the conditional probability of falling in the disadvantageous SEP category at each time period. To account for the case-control design, each exposure model was weighted by sampling fraction. The weights were stabilized by the marginal probability of falling in the disadvantageous SEP category at each time period. Once the stabilized inverse probability weights were computed, they were further

combined with time dependent sampling weights to account for the case-control design.[21] Sampling weights were defined as:

$$\text{Sampling weight} = [(1-\prod)/\prod] * \text{ncases/ncontrols},$$

where $\prod$ is the annual prevalence of oral cancer in India during the four-years of study, and ncases and ncontrols are the number of cases and controls in our sample. Finally, unadjusted logistic regression marginal structural outcome models were fit for each life-course model. In general, the outcome model took the form:

$$\text{Logit } \{Pr[Yg(A)=1]\} = \propto + \beta 1\ g(A)$$

where g(A) is a function of exposure, SEP, specific to each model. Additional technical details including those on exposure models and characteristics of stabilized weights are provided in Supplemental Appendix file, eTables 6 - 8.

We also fit a saturated all-trajectories model in which the other three models are nested. This model contained eight possible trajectories formed by binary SEP exposures measured over three periods of life.

Thirty-seven participants (17 controls and 20 cases) had missing values related to the main exposure. Therefore, we present our results on complete case analysis of 684 participants. Analyses were performed using Stata, version 13 SE (StataCorp. 2013, College Station, TX: StataCorp LP.). Annotated Stata codes are provided in Supplemental Appendix file, eTable 8.

3. RESULTS

Table 1 shows socio-demographic characteristics and measured potential confounders among cases and controls. The participants' age ranged from 32 to 88 years (mean=61 years) and the majority of the cases had a low level of education (78% of cases vs 50% of controls). The help of a proxy respondent was sought more rarely for controls (3%) than cases (14%). The majority of the participants belonged to the middle caste (81% of controls, 70% of cases). On an average, cases had a higher propensity for practicing all habits in each life period except for cigarette smoking. A higher proportion of cases than controls were exposed to disadvantageous SEP (60% vs 36% in childhood, 63% vs 35% in early adulthood, and 62% vs 33% in late adulthood).

Table 2 presents the association between life-course SEP and oral cancer under each conceptual model (IP weighted adjusted estimates. Crude estimates are presented in Supplemental Appendix file, eTable 9). Among the critical period models, being exposed to disadvantageous (vs. advantageous) SEP in childhood and early adulthood was associated with an increased odd of oral cancer (childhood: OR = 2.76, 95% CI: 1.99, 3.81; early adulthood: OR=1.84, 95% CI: 1.21, 2.79). In contrast, relative to an advantageous SEP, exposure to disadvantageous SEP in late adulthood was not associated with the disease (OR=0.92, 95% CI: 0.55, 1.54).

For the accumulation model, the odds of oral cancer increased with additional periods of socioeconomic disadvantage. Relative to never experiencing a period of disadvantageous SEP, experiencing one, two, and three periods of disadvantageous SEP yielded ORs of 2.56 (95% CI: 1.34, 4.87), 2.71 (95% CI: 1.44, 5.09), and 4.86 (95% CI: 2.61, 9.06), respectively.

Under the social mobility models, for childhood to early adulthood mobility, compared to stable advantageous SEP group, downward mobile (OR=2.75, 95% CI: 1.57, 4.83), upward mobile (OR=3.19, 95% CI: 1.83, 5.55) and stable disadvantageous (OR=4.06, 95% CI: 2.62, 6.28) trajectories were associated with increased odds for oral cancer.

The all-trajectories model (Table 2) showed that compared to non-exposure to disadvantageous SEP in all periods (0, 0, 0), the magnitude of ORs associated with trajectories in which individuals were exposed to disadvantageous SEP in childhood (1, 0, 0: OR= 4.37, 95% CI:1.83, 10.85 ; 1, 1, 0: OR=3.36, 95% CI 1.61, 6.99; 1, 0, 1: OR= 2.61, 95% CI: 1.16, 5.89; 1, 1, 1: OR=4.86, 95% CI: 2.61, 9.06) were larger than those of trajectories where participants were never exposed in childhood (0, 1, 0: OR=2.58, 95% CI: 1.15,5.80; 0, 0, 1: OR =1.00, 95% CI: 0.40, 2.53; 0, 1, 1: OR=2.25, 95% CI: 0.82, 6.21).

4. DISCUSSION

In this study, we examined the role of lifetime SEP on oral cancer risk comparing different life-course models and taking well-known behavioural risk factors into account under a time varying framework. Our findings indicate that an exposure to disadvantageous SEP in childhood may play a critical role in the development of oral cancer later in life.

Considered as the most fundamental of all life-course models,[22] the accumulation model implies cross-sectional clustering of (dis)advantages driven by social structure that accumulate longitudinally.[23] In our study, we found that the risk for oral cancer increased with the accumulation of disadvantageous SEP periods over the course of life. This finding is similar to the monotonically increasing risk pattern identified in life-course studies investigating other health outcomes.[24, 25] However, exploring other life-course models within this study provided further insight into this overall exposure-outcome relationship.

In line with studies investigating other chronic diseases including cancers,[24, 26, 27] our findings indicated that disadvantageous SEP during childhood and early adulthood increased the risk of oral cancer. The magnitude of association was higher for childhood. Interestingly, our findings from other models tested as well converged to indicate that childhood is a critical period for the risk of oral cancer. In our social mobility analyses, the magnitude of the OR associated with upward mobility from childhood to early adulthood was higher than that of downward mobility. This reflects the higher impact of disadvantageous SEP in childhood compared to the same exposure in early adulthood as observed from the critical period models. Also, the estimates from the all-trajectories model provided further evidence for the critical role an exposure to disadvantageous SEP in childhood may play in the increased risk for oral cancer later in life. A recent smaller study of 180 oral cancer cases and 272 controls from the nearby state of Karnataka, India, reported that a disadvantageous socioeconomic condition (measured using occupation of head of the household) in childhood had a significant effect on oral cancer risk that was not mediated through smoking, alcohol or paan chewing habits.[26]

The risk behaviours considered as time-varying variables in our study are usually considered to be affected by SEP and hence as mediators of the relationship between SEP and adult health outcomes. However, such behaviours (e.g., alcohol consumption) have been considered as determinants of socioeconomic consequences, especially in developing societies.[28] Although the state of Kerala ranks high in social development relative to other states, the state has one of the highest alcohol consumption levels in India.[29] Furthermore, alcohol consumption is highly correlated with tobacco habits. This strengthens our analytical approach considering the dual nature of these exposures as potential confounders and mediators.

The statistical evidence presented here does have biological plausibility. The adverse effects of an accumulation of socioeconomic disadvantage over an individual's life span can manifest biologically through increased allostatic load, impaired immune response, and specific genetic or epigenetic changes resulting in oral cancer.[9, 30] Of particular relevance to the critical period model, childhood represents a specific time of rapid development and vulnerability when exposures produce irreversible biological damage.[31] Childhood SEP captures different dimensions of adversity (e.g., poor nutrition) that may initiate the above carcinogenic processes in the oral cavity.[32]

There are several challenges in interpreting the results of our study. For example, although the results from the mobility models were in line with the gradient constraint hypothesis of social mobility,[33] the empirical difficulty in defining social mobility and associated life-course trajectories from limited time periods has been discussed in the literature.[24] In addition, although our sample size did exceed that of the majority of case-control studies exploring the SEP-oral cancer association,[7, 26] the results from social mobility and all-trajectories models tested were limited by the low numbers in some of the trajectories. There is also the potential for measurement error affecting our results. Our measure of SEP, an asset index,[17] may not have captured all aspects of SEP. However, asset indices serve as indicators of wealth and are particularly relevant to less industrialized societies.[16] Developing countries like India are more prone to high rates of short-term economic shock, and lack concrete socioeconomic classification systems such as those used in developed countries.[16, 17, 34] In addition, the cut-off points chosen to divide confounder vectors C2 into C2a and C2b (23 years) and C3 into C3a and C3b (50 years) were not based on statistical modelling, which might be a source of potential misclassification. However, we expect this to be negligible as our cut-off selection was based on the assumption that disadvantageous SEP during earlier stages of life (e.g., 17-23 years for early adulthood and 31-50 for late adulthood) is less likely to drive risk behaviours during these early stages (Figure 1). Moreover, behavioural factors in these earlier stages have higher probability to causally effect SEP later in life. Finally, recall bias is a well-recognized problem in case-control studies. Although, not a substitute for more reliable methods, we attempted to mitigate this bias by using a life-grid tool (in both cases and controls) that has been shown to improve recall.[35] We expect this to increase the expectation of non-differential misclassification of exposures, resulting in any bias towards the null. Relative measures of test-retest reliability for housing assets from this study are presented in Supplemental Appendix file, eTables5.

Several methodological strengths of our study also merit consideration. Cohort studies are not always feasible to investigate rare health outcomes. Our rigorous data collection procedures allowed us to analyse the temporal associations between exposures and potential confounders, as well as their time-varying nature. Leffondre et al,[21] developed weighted partial likelihood estimators for time-dependent exposures in a case-control setting. However, these estimators have not been extended to time-varying confounders. In this study, we employed a novel approach by combining these estimators with inverse probability weighting to account for both time-varying exposures and confounders. An additional complication in social epidemiology stems from the non-manipulable nature of social exposures such as SEP. We believe that, as there are numerous ways in which an individual may be "assigned" to a given level of SEP, each of which may have different impacts on the risk of oral cancer, interpreting associational estimates as causal effects is not possible.[36] However, our results do provide valid estimates of the socioeconomic distribution of oral cancer risk in Kerala, India. Furthermore, although the participation rate for controls were low in our study, a comparison of the housing assets of controls and data from the Census of India 2011, Kozhikode district, Kerala, showed that the distributions were similar (presented in Supplemental Appendix file, eTable 10) increasing the validity of our findings.[37]

4.1 Public health implications

Approximately 80% of oral cancer patients in India approaching health care facilities present with advanced disease, decreasing the success of treatments and overall survival rates.[38] The optimal solution to decrease the impact of these cancers on morbidity and mortality is to focus on comprehensive screening for these cancers (considered most amenable to early detection and treatment along with cervical and breast cancers), tailored to socioeconomic contexts in the population. Dentists, dental residents, nurses and dental hygienists can support such programmes by targeting populations with disadvantageous SEP. In its current form (e.g., use of toluidine blue dye, fluorescent imaging, brush biopsy), population-based screening for oral cancer is not cost-effective.[39] Furthermore, evidence on the effectiveness of the visual screening method is insufficient. However, the systematic examination of the oral cavity by dentists and physicians with particular attention to high-risk socioeconomic sub-groups (e.g., those exposed to a disadvantageous childhood SEP) is largely recommended. The coverage of such programmes can be increased through regular outreach programmes in dental schools (e.g., dental camps conducted by Government Dental College, Kozhikode, Kerala, in partnership with the National Service

Scheme of India). The opportunistic screening of high-risk group individuals and their referral to secondary prevention programs (e.g., alcohol and tobacco cessation) by medical service providers play a central role in a multi-disciplinary approach to the prevention of oral cancer.[39] Also, valid indicators of childhood SEP may be incorporated into oral cancer risk calculations and screening tools. Factoring in the negative effect of low SEP on oral cancer, specifically childhood SEP, can increase the precision of risk calculations and enhance the effectiveness of opportunistic screening.

5. CONCLUSIONS

To our knowledge, this is the first case-control study investigating the association between SEP over the life-course and oral cancer through the lens of multiple life-course models within a single study, and by considering the time-varying nature of SEP and multiple associated confounders over several periods of life. We used novel analytical approaches to existing methods adapted for a case-control study design to calculate our associational estimates. Multiple life-course models provided empirical evidence for the independent association of SEP during childhood with oral cancer risk in this Indian population. Addressing issues related to unfavourable social circumstances early in life may be beneficial in reducing the long-term burden of oral cancer in high-risk regions such as India.

Acknowledgements

We would like to thank our collaborators at the Indian site Dr I V Varghese, Dr Shameena Shiraz, and the HeNCe Life Study team, especially Dr Geneviève Castonguay for her valuable assistance in coordinating and monitoring the study. We would also like to thank Dr. Jay Kaufman, Professor, Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC, Canada, whose guidance helped in improving the manuscript.

Funding

This work was supported by the Canadian Institutes of Health Research [grant numbers: MOP 81172, MOP 111207]; Ministère du Développement économique, de l'Innovation et de l'Exportation du Québec: Programme de soutien à la recherche (PSR), volet: Soutien à des initiatives internationales de recherche et d'innovation (SIIRI). Belinda Nicolau holds a Canada Research Chair in Life Course Oral Epidemiology – Tier 2. Nicolas Schlecht is a Miriam Mandel Faculty Scholar in Cancer Research.

Conflicts of Interest: The authors declare that they have no conflict of interest.

Ethical approval: The study was approved by IRB and ethics committees of Government Dental and Medical colleges, Calicut, Kerala, India. All procedures performed in this study which involved human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Figure caption

Figure 1. Directed acyclic graph (DAG) representing the relationship between exposure, covariates and outcome, in the study from Kerala, India, 2008-2012 (n=684). **Oral cancer**: Outcome; **SEP:** Socioeconomic position, main exposure; **CH SEP**: SEP during childhood; **EAH SEP**: SEP during early adulthood; **LAH SEP**: SEP during late adulthood; **C0**: Vector representing baseline covariates, age, sex, caste i.e., hierarchy in Hindu religion (potential time-invariant confounders); **C1**: Vector representing education, health related behaviours (time-varying) of cigarette smoking, bidi smoking, paan chewing and alcohol consumption recorded during 0-16 years of age; **C2a**: Vector representing health related behaviours recorded during 17-23 years of age; **C2b**- Vector representing health related behaviours recorded during 24-30 years age; **C3a**- Vector representing health related behaviours recorded during 31-50 years of age; **C3b**- Vector representing health related behaviours recorded above 50 years

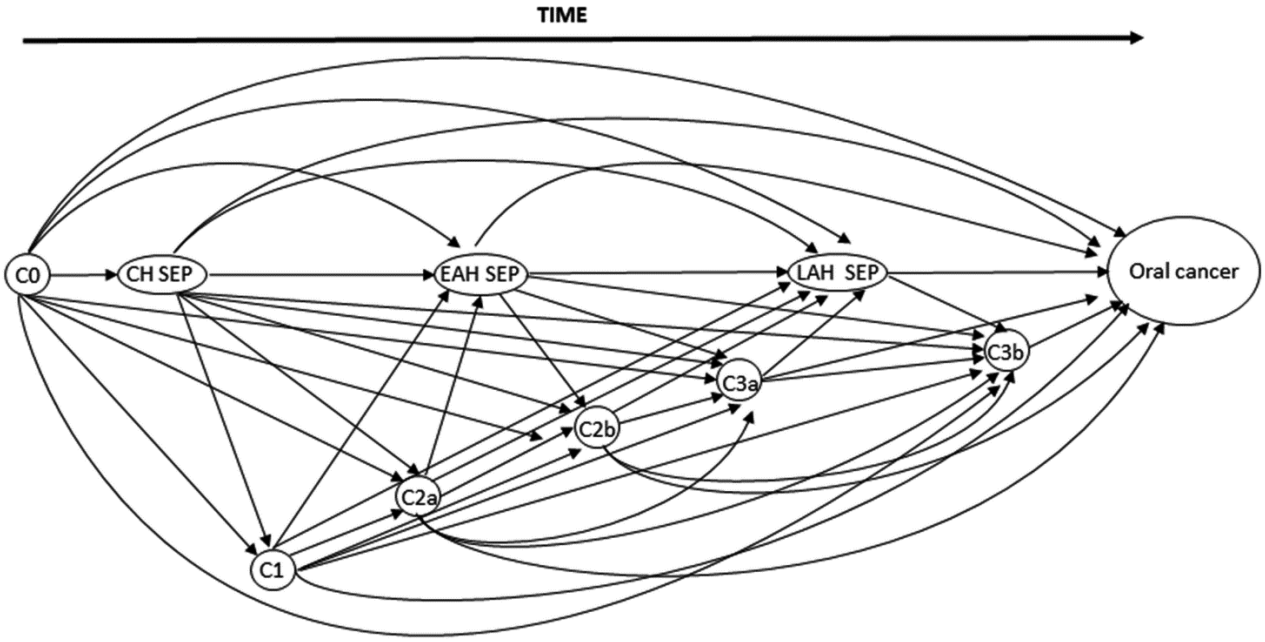

Table 1. Descriptive characteristics of oral cancer cases and controls from Kerala, India, 2008-12, (n=684)

	Controls (n=354)		Cases (n=330)	
	N (%)	**mean (SD)**	**N (%)**	**mean (SD)**
Age in years		61 (11)		61 (11)
Sex				
Female	163 (46)		149 (45)	
Male	191 (54)		181 (55)	
Education				
High	178 (50)		74 (22)	
Low	176 (50)		256 (78)	
Respondent type				
Use of proxy	11 (3)		46 (14)	
No use of proxy	343 (97)		284 (86)	
Caste				
Higher	51 (14)		26 (8)	
Middle	285 (81)		231 (70)	
Lower	18 (5)		73 (22)	
Time-varying risk behaviours				
During childhood (0-16 years)				
Cigarette smoking (pack-years)	25 (7)	0.08 (0.59)	13 (4)	0.05 (0.36)
Bidi smoking (pack-years)	45 (13)	0.14 (0.59)	66 (20)	0.23 (0.78)
Paan chewing (chew-years)	16 (5)	0.41 (3.50)	96 (29)	4.09 (9.29)
Alcohol consumption (drinks per week)	3 (0.8)	0.17 (2.29)	10 (3)	0.38 (2.70)
During early adulthood (17-23 years)				
Cigarette smoking (pack-years)	67 (19)	0.57 (1.73)	35 (11)	0.36 (1.87)
Bidi smoking (pack-years)	75 (21)	0.59 (1.94)	108 (33)	0.88 (2.10)
Paan chewing (chew-years)	30 (8)	1.72 (6.80)	156 (47)	12.44 (18.39)
Alcohol consumption (drinks per week)	32 (9)	2.54 (12.52)	50 (15)	4.30 (15.15)
During early adulthood (24-30 years)				
Cigarette smoking (pack-years)	97 (27)	1.26 (3.46)	63 (19)	0.88 (3.26)
Bidi smoking (pack-years)	81 (23)	0.91 (2.73)	120 (36)	1.60 (3.44)
Paan chewing (chew-years)	42 (12)	3.51 (11.67)	207 (63)	22.27 (28.24)
Alcohol consumption (drinks per week)	48 (14)	5.52 (33.18)	77 (23)	11.17 (43.33)
During late adulthood (31-50 years)				
Cigarette smoking (pack-years)	104 (29)	5.56 (14.33)	83 (25)	3.32 (9.58)
Bidi smoking (pack-years)	65 (18)	2.20 (7.59)	109 (33)	4.45 (9.19)
Paan chewing (chew-years)	57 (16)	15.76 (49.06)	237 (72)	94.95 (94.06)
Alcohol consumption (drinks per week)	59 (17)	6.65 (40.27)	90 (27)	15.37 (48.38)

Table 1 *Continued …*

	Controls (n=354)		Cases (n=330)	
	N (%)	**mean (SD)**	**N (%)**	**mean (SD)**
During late adulthood (51 years & above)				
Cigarette smoking (pack-years)	71 (20)	2.44 (9.77)	59 (18)	1.56 (5.34)
Bidi smoking (pack-years)	31 (9)	0.57 (2.63)	74 (22)	1.51 (4.24)
Paan chewing (chew-years)	52 (15)	15.76 (59.29)	183 (55)	60.41 (90.55)
Alcohol consumption (drinks per week)	42 (12)	2.47 (11.97)	60 (18)	11.71 (44.82)
SEP over the life-course				
Childhood SEP (0-16 years)				
Advantageous SEP	227 (64)		131 (40)	
Disadvantageous SEP	127 (36)		199 (60)	
Early adulthood SEP (17-30 years)				
Advantageous SEP	230 (65)		121 (37)	
Disadvantageous SEP	124 (35)		209 (63)	
Late adulthood SEP (above 30 years)				
Advantageous SEP	237 (67)		125 (38)	
Disadvantageous SEP	117 (33)		205 (62)	

SD: Standard deviation; SEP: Socioeconomic position.

Table 2. Odds ratios (adjusted for confounders using IP weighting) and 95% confidence intervals for risk of oral cancer under different life-course socioeconomic models in the study sample from Kerala, India, 2008-2012 (n=684)

Life-course SEP models	Levels of SEP (0 = Advantageous, 1= Disadvantageous)	Controls /Cases N	OR (95% CI)
Critical period models			
Childhood SEP	0[a]	227/131	Ref
	1	127/199	2.76 (1.99, 3.81)
Early adulthood SEP	0[a]	230/121	Ref
	1	124/209	1.84 (1.21, 2.79)
Late adulthood SEP	0[a]	237/125	Ref
	1	117/205	0.92 (0.55, 1.54)
Accumulation model			
Number of periods spent	0 periods[a]	162/53	Ref
in disadvantageous SEP	1 period	71/63	2.56 (1.34, 4.87)
over the life course	2 periods	66/92	2.71 (1.44, 5.09)
	3 periods	55/122	4.86 (2.61, 9.06)
Social mobility models			
Childhood-early adulthood SEP			
Stable advantageous	0,0[a]	190/79	Ref
Upward mobility	1, 0	40/42	3.19 (1.83, 5.55)
Downward mobility	0, 1	37/52	2.75 (1.57, 4.83)
Stable disadvantageous	1,1	87/157	4.06 (2.62, 6.28)
Early adulthood-late adulthood SEP			
Stable advantageous	0,0[a]	183/71	Ref
Upward mobility	1, 0	54/54	1.52 (0.80,2.87)
Downward mobility	0, 1	47/50	0.81 (0.40, 1.62)
Stable disadvantageous	1,1	70/155	1.53 (0.68, 3.41)
Saturated all-trajectories	0, 0, 0[a]	162/53	Ref
model[b]	1, 0, 0	21/18	4.37 (1.83,10.85)
(All SEP trajectories	0, 1, 0	22/19	2.58 (1.15,5.80)
across 3 life periods)	0, 0, 1	28/26	1.00 (0.40,2.53)
	1, 1, 0	32/35	3.36 (1.61,6.99)
	1, 0, 1	19/24	2.61 (1.16,5.89)
	0, 1, 1	15/33	2.25 (0.82,6.21)
	1, 1, 1	55/122	4.86 (2.61, 9.06)

SEP: socioeconomic position

[a] Reference category/ level within each SEP variable representing the specific life-course model.

[b] Categories/levels in the saturated all-trajectories model variable represents all possible 8 trajectories created from each binary SEP measure representing the three time periods.

References

1. Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. Oral Oncol 2009;45: 309-316.
2. Ferlay J, Soerjomataram I, Ervik M, et al. Globocan 2012 v1.0, cancer incidence and mortality worldwide: Iarc cancerbase no. 11 [internet]. Lyon, france: International agency for research on cancer; 2013. Available from: Http://globocan.Iarc.Fr, accessed on january 05, 2015., 2013.
3. Ferlay J, Shin H-R, Bray F, et al. Estimates of worldwide burden of cancer in 2008: Globocan 2008. Int J Cancer 2010;127: 2893-2917.
4. McKay AJ, Patel RKK, Majeed A. Strategies for tobacco control in india: A systematic review. PLoS One 2015;10: e0122610.
5. Petti S. Lifestyle risk factors for oral cancer. Oral Oncol 2009;45: 340-350.
6. Petersen PE. Strengthening the prevention of oral cancer: The who perspective. Community Dent Oral Epidemiol 2005;33: 397-399.
7. Conway DI, Petticrew M, Marlborough H, et al. Socioeconomic inequalities and oral cancer risk: A systematic review and meta-analysis of case-control studies. Int J Cancer 2008;122: 2811-2819.
8. Stringhini S, Sabia S, Shipley M, et al. Association of socioeconomic position with health behaviors and mortality. JAMA 2010;303: 1159-1166.
9. Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives. Int J Epidemiol 2002;31: 285-293.
10. VanderWeele TJ, Jackson JW, Li S. Causal inference and longitudinal data: A case study of religion and mental health. Soc Psychiatry Psychiatr Epidemiol 2016: 1-10.
11. Hallqvist J, Lynch J, Bartley M, et al. Can we disentangle life course processes of accumulation, critical period and social mobility? An analysis of disadvantaged socio-economic positions and myocardial infarction in the stockholm heart epidemiology program. Soc Sci Med 2004;58: 1555-1562.
12. Laprise C, Shahul HP, Madathil SA, et al. Periodontal diseases and risk of oral cancer in southern india: Results from the hence life study. Int J Cancer 2016;139: 1512-1519.
13. Madathil SA, Rousseau MC, Wynant W, et al. Nonlinear association between betel quid chewing and oral cancer: Implications for prevention. Oral Oncol 2016;60: 25-31.
14. Nishimoto, Pintos J, Schlecht NF, et al. Assessment of control selection bias in a hospital-based case-control study of upper aero-digestive tract cancers. J Cancer Epidemiol Prev 2002;7: 131-141.
15. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data-or tears: An application to educational enrollments in states of india. Demography 2001;38: 115-132.
16. Howe LD, Hargreaves JR, Gabrysch S, et al. Is the wealth index a proxy for consumption expenditure? A systematic review. J Epidemiol Community Health 2009;63: 871-877.
17. Gwatkin D, Rutstein S, Johnson K, et al. Socioeconomic differences in health, nutrition and population. Health, nutrition and population discussion paper, washington (d.C.): The world bank. 2000.
18. Akg center for research and studies, communist party of india (marxist), state committee, kerala. Education bill. Kerala,India: AKG Center for Research and Studies, 2009.

19. Robins JM, Hernán Ma, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology (Cambridge, Mass) 2000;11: 550-560.

20. Langholz B. Use of cohort information in the design and analysis of case-control studies. Scandinavian Journal of Statistics 2007;34: 120-136.

21. Leffondre K, Wynant W, Cao Z, et al. A weighted cox model for modelling time-dependent exposures in the analysis of case-control studies. Stat Med 2010;29: 839-850.

22. Blane D, Netuveli G, Stone J. The development of life course epidemiology. Rev Epidemiol Sante Publique 2007;55: 31-38.

23. Blane D. Social determinants of health--socioeconomic status, social class, and ethnicity. Am J Public Health 1995;85: 903-905.

24. Pollitt R, Rose K, Kaufman J. Evaluating the evidence for models of life course socioeconomic factors and cardiovascular outcomes: A systematic review. BMC Public Health 2005;5: 7.

25. Bernabe E, Suominen AL, Nordblad A, et al. Education level and oral health in finnish adults: Evidence from different lifecourse models. J Clin Periodontol 2011;38: 25-32; Peres MA, Peres KG, Thomson WM, et al. The influence of family income trajectories from birth to adulthood on adult oral health: Findings from the 1982 pelotas birth cohort. Am J Public Health 2011;101: 730-736.

26. Krishna Rao S, Mejia GC, Roberts-Thomson K, et al. Estimating the effect of childhood socioeconomic disadvantage on oral cancer in india using marginal structural models. Epidemiology 2015;26: 509-517.

27. Nicolau B, Netuveli G, Kim JW, et al. A life-course approach to assess psychosocial factors and periodontal disease. J Clin Periodontol 2007;34: 844-850; Vohra J, Marmot MG, Bauld L, et al. Socioeconomic position in childhood and cancer in adulthood: A rapid-review. J Epidemiol Community Health 2016;70: 629-634.

28. Global status report on alcohol and health, world health organisation. Who:Management of substance abuse, 2014. Available form: Http://www.Who.Int/substance_abuse/publications/global_alcohol_report/en/. 2014.

29. Kerala's 2005 human development report, united nations development program (undp). Available from: Http://www.In.Undp.Org/content/india/en/home/library/hdr/human-development reports/state_human_development_reports/kerala.Html, accessed on june 2nd, 2016. 2005; National sample survey office: Household consumption of various good and services in india 2011-2012, nss 68th round, ministry of statistics and programme implementation, government of india, june 2014. Available from: Http://mospi.Nic.In/mospi_new/upload/report_no558_rou68_30june14.Pdf. In: Implementation MoSaP, editor. India: Government of India, 2014.

30. McEwen BS. Stress, adaptation, and disease. Allostasis and allostatic load. Ann N Y Acad Sci 1998;840: 33-44; Stringhini S, Polidoro S, Sacerdote C, et al. Life-course socioeconomic status and DNA methylation of genes regulating inflammation. Int J Epidemiol 2015;44: 1320-1330.

31. Barker DJ. The fetal and infant origins of adult disease. BMJ : British Medical Journal 1990;301: 1111-1111.

32. Fagundes CP, Way B. Early-life stress and adult inflammation. Curr Dir Psychol Sci 2014;23: 277-283; Borghol N, Suderman M, McArdle W, et al. Associations with early-life socio-economic position in adult DNA methylation. Int J Epidemiol 2012;41: 62-74.

33. Blane D, Harding S, Rosato M. Does social mobility affect the size of the socioeconomic mortality differential?: Evidence from the office for national statistics longitudinal study. J R Stat Soc Ser A Stat Soc 1999;162: 59-70.

34. Galobardes B, Shaw M, Lawlor DA, et al. Indicators of socioeconomic position (part 1). J Epidemiol Community Health 2006;60: 7-12.

35. Berney L, Blane DB, Soc Sci M, et al. Collecting retrospective data: Accuracy of recall after 50 years judged against historical records. Soc Sci Med 1997;45: 1519-1525.

36. Naimi A, Kaufman J. Counterfactual theory in social epidemiology: Reconciling analysis and action for the social determinants of health. Current Epidemiology Reports 2015;2: 52-60.

37. Inda, 2011 census data. India: Office of the Registrar General & Census Commissioner, India, 2011 http://www.censusindia.gov.in/2011census/dchb/KerlaA.html.

38. Kulkarni MR. Head and neck cancer burden in india. Int J Head Neck Surg 2013;4: 29-35.

39. Kujan O, Glenny AM, Duxbury J, et al. Evaluation of screening strategies for improving oral cancer mortality: A cochrane systematic review. J Dent Educ 2005;69: 255-265.

Differential impact of socioeconomic position across life on oral cancer risk in Kerala, India: An investigation of life-course models under a time-varying framework

eAppendix

Measurement of socioeconomic position (SEP)

Asset/wealth index was created from a list of questions on various assets (housing characteristics, durable assets and access to services) available at the participant's longest place of residence during three time periods: childhood (0-16 years), early adulthood (17-30 years), and late adulthood (above 30 years). As given in Appendix Table 1, information on nine assets/items from childhood, eleven from early adulthood and twelve from late adulthood were used. The nominal responses to each of these questions were binary coded based on type of material used and facilities available, contextual to Kerala, India. A tetrachoric correlation matrix (Debelak and Tran 2013) was created from these binary variables for each life period (Appendix Tables 2,3,4). If any variable correlated highly (|0.8|) with other variables, only one variable from the group of correlated variables were retained for further analysis. In addition, variables were excluded in stepwise manner until a factorable correlation matrix with Kaiser-Meyer-Olkin (KMO) value > 0.7 was attained for each period separately (Balen et al. 2010). Assets with low test-retest reliability (inter class correlation) were also removed (Appendix Table 5). Final variables retained in the matrix for each period were; Childhood: crowding, floor, wall, window, water, bath, clock, KMO=0.832; Early adulthood: crowding, wall, window, water, clock, bicycle; KMO=0.771; Late

adulthood: Crowding, wall, window, water, clock, radio, television, phone, KMO=0.801. A principal component analysis was conducted without rotation on the final correlation matrices to assess dimensionality of the assets, and the first component that explained maximum variance in each life period (childhood 1st component explained 65% of variance, 64% each for early and late adulthood) was extracted (Filmer and Pritchett 2001). Scores were predicted out of these components. Each of the continues score for each life period was then dichotomized using the median of the distribution as cut-off generating respective binary variable representing SEP (0= advantageous SEP, 1= disadvantageous SEP) for childhood, early and late adulthood.

SEP exposure measure for critical period models

The binary variable (0-advantageous SEP, 1-disadvantageous SEP) representing SEP in childhood, early, and late adulthood were used as the main exposure in the critical period model representing each of these life periods.

SEP exposure measure for accumulation model

A summation of the binary variables representing SEP in each life period generated a variable with four categories with increasing periods of exposure to disadvantageous SEP. This variable represented the accumulation model. The variable was coded as: 0=0 period– participants who were in advantageous SEP in all 3 periods of life; 1=1 period-participants who were exposed to disadvantageous SEP in any 1 period and non-exposed in any 2 periods of life; 2=2 periods - participants who were exposed to disadvantageous SEP in any 2 periods and non-exposed in any 1 period of life; and 3= 3 periods-participants who were exposed to disadvantageous SEP in all three periods of life.

SEP exposure measure for social mobility models

Two models were tested for mobility; childhood to early adulthood mobility, and early to late adulthood mobility.

Childhood to early adulthood mobility - The SEP measure representing this model was a 4-category variable. *Stable advantageous SEP (0, 0)*: Participants who maintained a stable advantageous SEP in both childhood and early adulthood irrespective of their SEP in late adulthood, were coded as 0. *Upward mobility (1, 0)*: Participants who were exposed to a disadvantageous SEP in childhood but went on to attain an advantageous SEP in early adulthood irrespective of their SEP in late adulthood were coded as 1. *Downward mobility (0, 1)*: Participants who had an advantageous SEP in childhood but disadvantageous SEP in early adulthood irrespective of their SEP in late adulthood were coded as 2. *Stable disadvantageous SEP (1, 1)*: Participants who maintained a stable disadvantageous SEP in both childhood and early adulthood irrespective of their SEP in late adulthood, were coded as 3;

Early to late adulthood mobility - A similar strategy was adopted to create the 4 category SEP variable representing social mobility between early and late adulthood by considering participants' SEP in these 2 periods of life.

References

Balen J, McManus DP, Li Y-S, Zhao Z-Y, Yuan L-P, Utzinger J, Williams GM, Li Y, Ren M-Y, Liu Z-C et al. 2010. Comparison of two approaches for measuring household wealth via an asset-based index in rural and peri-urban settings of hunan province, china. Emerging Themes in Epidemiology. 7:7-7.

Debelak R, Tran US. 2013. Principal component analysis of smoothed tetrachoric correlation matrices as a measure of dimensionality. Educational and Psychological Measurement. 73(1):63-77.

Filmer D, Pritchett LH. 2001. Estimating wealth effects without expenditure data-or tears: An application to educational enrollments in states of india. Demography. 38(1):115-132.

Additional details on confounders

Categorization of Education

Detailed information regarding education was collected from each participant in our study. We used number of years of formal education in the form of a binary variable (0: high education; 1: low education) as an indicator. However, the measure of education is subjected to bias if the differences in birth cohorts of participants from a range of age groups included in a study are unaccounted for. With respect to the Kerala study site, considerable educational and sociopolitical reforms took place in the mid1950s, which changed the landscape of education in this state of India This information was used to mitigate bias in the categorization of education. The participants were first divided into 2 groups: older: those born before 1950, younger: those born after 1950). For the older cohort, 0-3 years of formal education was considered low level, and 4 years and above was considered as high level of education. For the younger cohort, 8 years of formal education as used as the cut-off for this binary categorization.

Caste

Caste refers to hereditary classes or hierarchy in Hindu society in India, based on occupation. Caste is a concept confined to India. The categories in caste can be specific to states in India. In this study, information on the caste was collected on each participant (1: Forward caste, 2: backward caste, 3: other backward caste, 4: scheduled caste, 5: scheduled tribe, 6: None of the above, : NA/Christians) based on the list of castes determined by the Government of Kerala, India. Using this information, the caste variable was categorized into (0=higher caste, 1=middle caste comprising of backward caste, 2=other backward/scheduled caste/scheduled tribe/others).

Tobacco-pack years: It is defined as the product of number of packs (smoked per day and duration of smoking.

Standard drinks: There is no consensus on the definition of a standard drink in India (13 to 28g of pure ethanol). Thus, in this study, we divided milliliters of ethanol consumed per week by 18 (standard drink=18ml of alcohol containing 14g of pure ethanol) to make it equivalent and comparable to North American standards.

Temporal relationship of confounders in relation to SEP in three periods of life and oral cancer

The temporal ordering of exposures and covariates with respect to the outcome is imperative when testing life-course model. Furthermore, to estimate causal effects (or when applying frameworks for causal inference or associated analytical techniques), the precedence of the causal factor in relation to its effect, is of absolute necessity. Whereas temporal ordering is easier in studies capturing longitudinal data, it is a challenge in case-control studies. But our detailed and comprehensive data collection methods. and techniques to handle the details on confounders in our life-course based study allowed us to achieve an approximate temporal ordering of variables with respect to SEP in several periods of life and oral cancer diagnosis. As shown in **the Figure** in the manuscript (causal graph), the vector *C0* represented age of the participant, and time-invariant covariates such as sex and caste that temporally precede every other variable under consideration. The vector *C1* represented covariates that were measured for the period between 0-16 years of age. We included education in *C1* because it is usually attained during this period, and could causally affect the subsequent life events of an individual. Other variables represented in *C1* and subsequent vectors *C2a, C2b, C3a* and *C3b* were time-varying risk behaviours (cigarette, bidi, paan and alcohol use). The cumulative measures of these risk behaviours were calculated for 0-16 years, 17-23 years, 24-30 years, 31 -50 years, and above 50 years. Risk factors collected for the period between 0-16 years might be an effect rather than cause of SEP between 0-16 years of age and were included in *C1*. However, we suspected that the association between *early adulthood SEP* (17-30 years) and habits captured during 17-30 years, was bi-directional, that is, SEP and habits can influence each other causally. Bidirectional arrows cannot occur in causal structures at the same time point. To overcome this, we split the habits in this period into vectors C2a (17-23 years) and C2b (24-30 years). This was done assuming that *C2a* would be affected by *C0, C1* and *CH SEP*, but would influence part of SEP in 17-30 years and other subsequent variables. And *C2b* would be affected by *C0, C1, C2a, CH SEP* and *EAH SEP*. The choice of cut-point (i.e., 23 years) was arbitrary. A similar strategy was used with risk behaviours recorded for above 30 years of age. Risk behaviours recorded during the period 31-50 years of age were represented by C3a, and those recorded above for 50 years (the eldest participant was 88 years old) were represented by C3b. This approximate temporal ordering identified complex feed-back loops between the variables under study as any given variable/vector represented in **Figure** had an arrow pointing from them to any other variable/vector temporally subsequent to it.

eTable1: List of housing assets/items, their categories, corresponding binary Stata codes and their proportion, if selected for creating the SEP measure for childhood, early or late adulthood.

Assets/items	Categories (Stata code)	Proportion, if used in childhood (%)	Proportion, if used in early adulthood	Proportion, if used in late adulthood
Crowding	Absent (0)	45.61	41.81	50.58
	Present (1)	54.39	58.19	49.42
Material of floor used	High cost (0)	85.23	65.50	17.69
	Low cost (1)	14.77	34.50	82.31
Material of roof used	High cost 0)	68.57	50.44	12.72
	Low cost (1)	31.43	49.56	87.28
Material of wall used	High cost (0)	77.92	65.06	23.54
	Low cost (1)	22.08	34.94	76.46
Windows	High cost (0)	34.06	19.15	06.14
	Low cost (1)	65.94	80.85	93.86
Water source	Protected (0)	46.35	34.50	09.06
	Unprotected (1)	53.65	65.50	90.94
Bathroom	Present (0)	79.39	51.46	06.29
	Absent (1)	20.61	48.54	93.71
Clock	Present (0)	84.06	58.19	10.67
	Absent (1)	15.94	41.81	89.33
Radio	Present (0)	91.08	73.39	31.58
	Absent (1)	08.92	26.61	68.42
Bicycle	Present (0)		90.64	
	Absent (1)		09.36	
Electricity	Present (0)		75.15	19.30
	Absent (1)		24.85	80.70
Television	Present (0)			42.11
	Absent (1)			57.89
Phone	Present (0)			32.31
	Absent (1)			67.69

0 represents advantageous SEP, and 1 represents disadvantageous SEP

eTable2: Tetrachoric correlation matrix for items recorded in childhood

	CH_crowd	CH_floor	CH_roof	CH_wall	CH_wind	CH_water	CH_bath	CH_clock	CH_radio
CH_crowd	1.0000								
CH_floor	0.4674	1.0000							
CH_roof	0.5912	0.8038	1.0000						
CH_wall	0.5362	0.7876	0.8618	1.0000					
CH_wind	0.4474	0.6791	0.7352	0.6613	1.0000				
CH_water	0.4203	0.5282	0.5891	0.5981	0.5493	1.0000			
CH_bath	0.4827	0.7544	0.7556	0.6522	0.4896	0.5396	1.0000		
CH_clock	0.5790	0.7576	0.7432	0.7788	0.4623	0.4433	0.7568	1.0000	
CH_radio	0.5581	0.7296	0.7272	0.7147	0.5562	0.5295	0.7673	0.9068	1.0000

eTable3: Tetrachoric correlation matrix for items recorded in early adulthood

	EAH_crowd	EAH_floor	EAH_roof	EAH_wall	EAH_wind	EAH_water	EAH_bath	EAH_elect	EAH_clock	EAH_radio	EAH_cycle
EAH_crowd	1.0000										
EAH_floor	0.5091	1.0000									
EAH_roof	0.5045	0.8311	1.0000								
EAH_wall	0.4791	0.8464	0.8302	1.0000							
EAH_wind	0.3798	0.6649	0.6962	0.8196	1.0000						
EAH_water	0.3618	0.6284	0.6184	0.6894	0.5987	1.0000					
EAH_bath	0.4455	0.7554	0.7512	0.6827	0.5469	0.6093	1.0000				
EAH_elect	0.4269	0.7462	0.7441	0.7474	0.5224	0.5759	0.8046	1.0000			
EAH_clock	0.3843	0.6544	0.6289	0.6650	0.3359	0.4505	0.6566	0.8448	1.0000		
EAH_radio	0.4532	0.7263	0.6860	0.6790	0.5206	0.5628	0.8108	0.8114	0.8704	1.0000	
EAH_cycle	0.3946	0.5108	0.5326	0.5448	0.3879	0.6253	0.5962	0.5736	0.7142	0.8102	1.0000

eTable4: Tetrachoric correlation matrix for items recorded in late adulthood

	LAH_crowd	LAH_floor	LAH_roof	LAH_wall	LAH_wind	LAH_water	LAH_bath	LAH_clock	LAH_radio	LAH_elect	LAH_tv	LAH_phone
LAH_crowd	1.0000											
LAH_floor	0.3263	1.0000										
LAH_roof	0.3178	0.8622	1.0000									
LAH_wall	0.3237	0.8811	0.8672	1.0000								
LAH_wind	0.2523	0.5743	0.6523	0.5789	1.0000							
LAH_water	0.2568	0.4108	0.4918	0.4424	0.4123	1.0000						
LAH_bath	0.2943	0.7639	0.7493	0.7337	0.5949	0.5375	1.0000					
LAH_clock	0.1781	0.6312	0.5693	0.6312	0.3599	0.3192	0.7153	1.0000				
LAH_radio	0.3373	0.4644	0.5405	0.4725	0.3582	0.3729	0.5895	0.7428	1.0000			
LAH_elect	0.3161	0.7030	0.7312	0.6030	0.5455	0.4411	0.8708	0.6296	0.5432	1.0000		
LAH_tv	0.3371	0.7417	0.6730	0.6848	0.5621	0.4188	1.0000	0.7670	0.5826	0.8759	1.0000	
LAH_phone	0.2706	0.5839	0.6039	0.5992	0.5165	0.4240	1.0000	0.7521	0.5584	0.7088	0.8120	1.0000

eTable5: Relative measures of test-retest reliability for housing-based assets used to create SEP measures for childhood, early and late adulthood periods.

Assets/items	N	Childhood			Early adulthood			Late adulthood		
		Pearson correlation	Intra class correlation	95% confidence interval	Pearson correlation	Intra class correlation	95% confidence interval	Pearson correlation	Intra class correlation	95% confidence interval
Crowding	46[a]	0.91	0.95	0.92, 0.98	0.86	0.93	0.87, 0.96	0.74	0.85	0.73, 0.92
Material of floor	46[a]	0.91	0.95	0.92, 0.98	0.99	0.99	0.99, 0.99	0.63	0.79	0.58, 0.87
Material of roof	46[a]	0.99	0.99	0.99, 0.99	0.99	0.99	0.99, 0.99	0.99	0.99	0.99, 0.99
Material of wall	46[a]	0.99	0.99	0.99, 0.99	0.95	0.98	0.96, 0.99	0.92	0.96	0.93, 0.98
Windows	46[a]	0.86	0.93	0.86, 0.56	0.99	0.99	0.99, 0.99	0.99	0.99	0.99, 0.99
Water source	46[a]	0.92	0.96	0.92, 0.98	0.90	0.95	0.90, 0.97	0.85	0.98	0.94, 0.99
Bathroom	46[a]	0.99	0.99	0.98, 0.99	0.87	0.93	0.87, 0.96	0.70	0.80	0.63, 0.89
Clock	46[a]	0.83	0.76	0.64, 0.95	0.82	0.79	0.55, 0.94	0.84	0.77	0.59, 0.96
Radio	46[a]	0.75	0.85	0.72, 0.91	0.79	0.87	0.76, 0.93	0.69	0.79	0.62, 0.88
Bicycle	46[a]				0.73	0.82	0.67, 0.90	0.64	0.80	0.62, 0.88
Electricity	46[a]				0.92	0.97	0.95, 0.98	0.91	0.95	0.92, 0.96
Television	46[a]							0.85	0.92	0.85, 0.54
Phone	46[a]							0.87	0.93	0.87, 0.96

[a] Among the sample of 721 participants recruited in total at the Indian site, re-interviews were conducted for 46 randomly selected participants, 6 to 12 weeks after the original interview. The above measures were estimated among these participants.

eTable6: Conceptual life course models, corresponding trajectories, causal contrasts and regression models

Conceptual Model	Levels of exposure (0=No, 1=Yes)	Contrast for each trajectory	Marginal structural regression models
All-trajectories saturated model			
Never exposed	0, 0, 0		
Exposed in CH (A_{100}) vs never exposed	1, 0, 0	$E[Y_{100}- Y_{000}]$	logit $\{Pr[Y_g(SEP)]\} = \alpha + \beta_1 g(SEP)$
Exposed in EAH (A_{010}) vs never exposed	0, 1, 0	$E[Y_{010}- Y_{000}]$	
Exposed in LAH (A_{001}) vs never exposed	0, 0, 1	$E[Y_{001}- Y_{000}]$	g(SEP)=function of 8 category variable
Exposed in CH & EAH (A_{110}) vs never	1, 1, 0	$E[Y_{110}- Y_{000}]$	involving all 8 life-course trajectories.
Exposed in CH & LAH (A_{101}) vs never	1, 0, 1	$E[Y_{101}- Y_{000}]$	Betas correspond to estimates for
Exposed in EAH & LAH (A_{011}) vs never	0, 1, 1	$E[Y_{011}- Y_{000}]$	each contrast in column 3.
Exposed in CH, EAH&LAH (A_{111}) vs never	1, 1, 1	$E[Y_{111}- Y_{000}]$	
Accumulation model			
Never exposed	0 periods		
Exposed at 1 time point vs never exposed	1 period	$E(Y_{100,010,001}-Y_{000})$	logit $\{Pr[Y_g(SEP)]\} = \alpha + \beta_1 g(SEP)$
Exposed at 2 time points vs never	2 periods	$E(Y_{110,101,011}-Y_{000})$	g(SEP)= function of 4 category variable
Exposed at 3 time points vs never	3 periods	$E(Y_{111}-Y_{000})$	involving specific combination of
Critical period			
Exposed in CH vs unexposed in CH	1 vs 0	$E(Y_{1**}-Y_{0**})$[a]	logit $\{Pr[Y(csep)=1]\} = \alpha + \beta_1 *g(csep)$
Exposed in EAH vs unexposed in EAH	1 vs 0	$E(Y_{*1*}-Y_{*0*})$ [a]	logit $\{Pr[Y(esep)=1]\} = \alpha + \beta_1 *g(esep)$
Exposed in LAH vs unexposed in LAH	1 vs 0	$E(Y_{**1}-Y_{**0})$ [a]	logit $\{Pr[Y(lsep)=1]\} = \alpha + \beta_1 * g(lsep)$
Social mobility model			
Childhood to early adulthood			logit $\{Pr[Y_g(SEP)]\} =\alpha +\beta 1\ g(SEP)$
Stable advantageous	0,0		g(SEP)= function of 4 category variable
Upward mobility	1, 0	$E[Y_{10*}-Y_{00*}]$ [a]	involving specific combination of
Downward mobility	0, 1	$E[Y_{01*}-Y_{00*]}$ [a]	trajectories over CH and EAH SEP
Stable disadvantageous	1, 1	$E[Y_{11*}-Y_{00*]}$ [a]	
Early to Late adulthood			
Stable advantageous	0,0		logit $\{Pr[Y_g(SEP)]\} =\alpha +\beta 1\ g(SEP)$
Upward mobility	1, 0	$E[Y_{*10}-Y_{*00]}$ [a]	g(SEP)= function of 4 category variable
Downward mobility	0, 1	$E[Y_{*01}-Y_{*00]}$ [a]	involving specific combination of
Stable disadvantageous	1,1	$E[Y_{*11}-Y_{*00]}$ [a]	trajectories over EAH and LAH SEP

Abbreviations: CH=childhood; EAH= early adulthood; LAH= late adulthood; SEP= Socioeconomic position; csep= childhood SEP; esep= early adulthood SEP; lsep= late adulthood SEP
If A is the exposure level, A=1 would represent exposed to disadvantageous SEP, and A=0 would be non-exposure.
[a] *can take any value between 0 and

eTable7: Summary statistics for minimally stabilized inverse probability weights based on which the final weights for the outcome marginal structural models were created.

Stabilized inverse probability weights	N	Mean	SD	Min	Max
W1	684	1.02	0.41	0.42	4.10
W2	684	1.01	0.84	0.26	5.66
W3	684	1.13	1.29	0.33	7.19
W12	684	1.03	0.90	0.16	5.97
W123	684	1.16	1.64	0.06	13.47

W1 – Stabilized inverse probability weight for childhood; W2 – Stabilized inverse probability weight for early adulthood; W3 –Stabilized inverse probability weight for late adulthood; W12- Product of W1 and W2; W123- Product of W1, W2 and W3. Please see eTable 8 for details

eTable8: Annotated **Stata codes** for exposure weights, description of SEP exposure weight specification, inverse-probability weights and outcome marginal structural models

```
// For creating inverse probability weight for childhood (W1)
logit csep [pw=Sampfrac] // [for PP of numerator]
logit csep age* sex caste [pw=Sampfrac] // (for PP of denominator)

// For creating inverse probability weight for early adulthood (W2)
logit esep [pw=Sampfrac] // (for numerator PP)
logit esep csep age* sex caste edu C1Cig* C1Bidi* C1Chew* C1Drink* C2aCig* C2aBidi* C2aChew* ///
C2aDrink* [pw=Sampfrac] // (for PP denominator)

// For creating inverse probability weight for late adulthood (W3)
logit lsep [pw=Sampfrac] // (for numerator PP)
logit lsep esep csep age* sex caste edu C1Cig* C1Bidi* C1Chew* C1Drink* C2aCig* C2aBidi* ///
C2aChew* C2aDrink* C2bCig* C2bBidi* C2bChew* C2bDrink* C3aCig* C3aBidi* C3aChew* C3aDrink* [pw=Sampfrac]
// (for PP denominator)

/*csep=childhood SEP, esep=early adulthood SEP, lsep= late adulthood SEP, edu = education, Cig=cigarette
smoking, Bidi=Bidi smoking, Chew=paan chewing, Drink= alcohol consumption, *= entered as spline variable,
PP=predicted probability; Sampfrac =sampling fraction, (VanderWeele & Vansteelandt 2010). Please refer to
Fig 1 to understand the confounder variables from their respective prefix (e.g., C1, C2a, C2b, C3a) */

// Weight multiplication [SampW=time dependent sampling weight (Leffondre et al, 2010]
sW1=W1*SampW
sW12= W12*SampW // where W12=W1*W2
sW123= W123*SampW // where W123=W1*W2*W3

// Outcome marginal structural models (unadjusted logistic regression models on the pseudo-population)
logistic Status i.Traj_SEP [pw=sW123 ] //  (saturated all-trajectories model)
logistic Status i.AccSEP [pw=sW123] //(accumulation model)
logistic Status csep [pw=sW1] // (childhood critical period)
logistic Status esep [pw=sW12] // (early adulthood critical period)
logistic Status lsep [pw=sW123] // (late adulthood critical period)
logistic Status i.ce_mob [pw=sW12] // ce_mob= childhood to early adulthood mobility SEP variable
logistic Status i.ea_mob [pw=sW123] // ea_mob= early to late adulthood mobility SEP variable
```

eTable 9. Odds ratios (unadjusted for behavioural risk factor confounders) and 95% confidence intervals for risk of oral cancer under different life-course socioeconomic models in the study sample from Kerala, India, 2008-2012 (n=684)

Life-course SEP models	Levels of SEP (0 = Advantageous, 1= Disadvantageous)	Controls /Cases N	OR (95% CI)
Critical period models			
Childhood SEP	0[a]	227/131	Ref
	1	127/199	2.71 (1.99, 3.70)
Early adulthood SEP	0[a]	230/121	Ref
	1	124/209	3.20 (2.34, 4.38)
Late adulthood SEP	0[a]	237/125	Ref
	1	117/205	3.32 (2.43, 4.54)
Accumulation model			
Number of periods spent	0 periods[a]	162/53	Ref
in disadvantageous SEP	1 period	71/63	2.71 (1.71, 4.29)
over the life course	2 periods	66/92	4.26 (2.73, 6.63)
	3 periods	55/122	6.78 (4.45, 10.57)
Social mobility models			
Childhood-early adulthood SEP			
Stable advantageous	0,0[a]	190/79	Ref
Upward mobility	1, 0	40/42	2.52 (1.52, 4.19)
Downward mobility	0, 1	37/52	3.38 (2.06, 5.55)
Stable disadvantageous	1,1	87/157	4.34 (2.99, 6.29)
Early adulthood-late adulthood SEP			
Stable advantageous	0,0[a]	183/71	Ref
Upward mobility	1, 0	54/54	2..58 (1.62,4.10)
Downward mobility	0, 1	47/50	2.74 (1.69, 4.44)
Stable disadvantageous	1,1	70/155	5.71 (3.85, 8.45)
Saturated all-trajectories	0, 0, 0[a]	162/53	Ref
model[b]	1, 0, 0	21/18	2.62 (1.30,5.28)
(All SEP trajectories	0, 1, 0	22/19	2.64 (1.33,5.25)
across 3 life periods)	0, 0, 1	28/26	2.84 (1.53,5.26)
	1, 1, 0	32/35	3.34 (1.88,5.92)
	1, 0, 1	19/24	3.86 (1.96,7.59)
	0, 1, 1	15/33	6.72 (3.39,13.33)
	1, 1, 1	55/122	6.78 (4.35, 10.57)

SEP: socioeconomic position

[a] Reference category/ level within each SEP variable representing the specific life-course model.

[b] Categories/levels in the saturated all-trajectories model variable represents all possible 8 trajectories created from each binary SEP measure representing the three time periods.

eTable10: A comparison of percentage distribution of housing assets (longest residence in late adulthood) of controls recruited in HeNCe life study Kozhikode, India site and available data from the Census of India 2011, Kozhikode district, Kerala.

Housing Assets	Census of India 2001 Calicut district, Kerala %*	HeNCe Life study, India site %*
Number of rooms		
1	1.1	2.0
2	7.0	9.0
3	28.0	29.0
3+	34.0	30.0
Water system		
Tap water	21.0	22.0
Well	72.8	74.0
Spring/River/Canal/Tank/Pond	1.8	1.4
Electricity		
Yes	93.8	94.0
No	6.2	6.0
Sanitation		
Septic tank/latrine/slab covered	83.0	90.0
others	17.0	10.0
Kitchen facility		
Yes	97.1	96.0
No	2.7	3.4
No self cooking	0.2	0.5
TV (Present)	71.76	73
Telephone (Present)	78	74
Scooter/motorbike (Present)	25	33
Car/Jeep (Present)	8	11

*Cumulative percentage may not add to 100%. The comparison is for assets whose information was available in both data sets.