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.