


BMJ Open Identifying the latent classes of modifiable risk behaviours among diabetic and hypertensive individuals in Northeastern India: a population-based cross-sectional study

Strong P Marbaniang ^{1,2}, Hemkothang Lhungdim,¹
Holendro Singh Chungkham^{3,4}

To cite: Marbaniang SP, Lhungdim H, Chungkham HS. Identifying the latent classes of modifiable risk behaviours among diabetic and hypertensive individuals in Northeastern India: a population-based cross-sectional study. *BMJ Open* 2022;**12**:e053757. doi:10.1136/bmjopen-2021-053757

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-053757>).

Received 25 May 2021
Accepted 04 February 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Public Health and Mortality Studies, International Institute for Population Sciences, Mumbai, India

²Department of Statistics, Sankardev College, Shillong, Meghalaya

³Applied and Official Statistics Unit, Indian Statistical Institute, North-East Centre at Tezpur, Tezpur, India

⁴Stress Research Institute, Stockholm University, Stockholm, Sweden

Correspondence to

Strong P Marbaniang;
strongmarbaniang@yahoo.com

ABSTRACT

Objective To identify the latent classes of modifiable risk factors among the patients with diabetes and hypertension based on the observed indicator variables: smoking, alcohol, aerated drinks, overweight or obesity, diabetes and hypertension. We hypothesised that the study population diagnosed with diabetes or hypertension is homogeneous with respect to the modifiable risk factors.

Design A cross-sectional study using a stratified random sampling method and a nationally representative large-scale survey.

Setting and participants Data come from the fourth round of the Indian National Family Health Survey, 2015–2016. Respondents aged 15–49 years who were diagnosed with either diabetes or hypertension or both were included. The total sample is 22 249, out of which 3284 were men and 18 965 were women.

Primary and secondary outcome measures The observed variables used as latent indicators are the following: smoking, alcohol, aerated drinks, overweight or obesity, diabetes and hypertension. The concomitant variables include age, gender, education, marital status and household wealth index. Latent class model was used to simultaneously identify the latent class and to determine the association between the concomitant variables and the latent classes.

Results Three latent classes were identified and labelled as class 1: 'diabetic with low-risk lifestyle' (21%), class 2: 'high-risk lifestyle' (8%) and class 3: 'hypertensive with low-risk lifestyle' (71%). Class 1 is characterised by those with a high probability of having diabetes and low probability of smoking and drinking alcohol. Class 2 is characterised by a high probability of smoking and drinking alcohol and class 3 by a high probability of having high blood pressure and low probability of smoking and drinking alcohol.

Conclusions Co-occurrence of smoking and alcohol consumption was prevalent in men, while excess body weight and high blood pressure were prevalent in women. Policy and programmes in Northeastern India should focus on targeting multiple modifiable risk behaviours that co-occur within an individual.

Strengths and limitations of this study

- The sample for the study is relatively large as it uses a nationally representative large-scale data.
- Respondents' status of diabetes and hypertension is diagnosed during fieldwork (on the day of interview) using a standard biomarker instrument and measurement.
- Latent class analysis is a person-centred technique and the appropriate method for explaining the clustering of health risk behaviours.
- The cross-sectional design of the study restricts us to measure the causal effect.
- Modifiable risk factors were collected based on dichotomous responses, which limit measuring the intensity of substance use.

INTRODUCTION

Effective lifestyle modification plans and affordable approaches to control high blood sugar and high blood pressure (BP) level benefit the at-risk population.^{1 2} The leading causes of mortality due to chronic diseases in adults, particularly older adults, are linked to unhealthy lifestyle and behaviours,³ such as tobacco consumption, physical inactivity, excess alcohol consumption and poor diet. An alarmingly high prevalence of hypertension in Northeastern (NE) India has been reported in recent studies.^{4–6} Sikkim has the highest prevalence of hypertension among all the Indian states, while other NE states such as Nagaland, Assam, Arunachal Pradesh, Mizoram and Tripura have prevalence above 11.0%, which is well above the national average.^{5 6} Further, according to the Global Burden of Disease study, the percentage change of diabetes prevalence between 1990 and 2016 has increased by more than 20% across the states of NE India.⁷ This percentage

change was highest in the states of Sikkim, Arunachal Pradesh and Nagaland.

Modifiable health risk behaviours such as smoking, heavy consumption of alcohol, low physical activity and unhealthy diet characterised by high intake of sugar and fats, low consumption of fruits and vegetables are the major causes of non-communicable diseases (NCDs)⁸ which are strongly linked with early mortality.^{9 10} Usually, these risk behaviours exist simultaneously or cluster within individuals.¹¹ It is argued that lifestyle risk factors within individuals are not random, but more likely to cluster with other unhealthy behaviours.^{12 13} Further, when two risk behaviours cluster together with each other more than the other factors, it may suggest that they could be influenced by a common source.¹⁴ Moreover, modifiable risk behaviours for chronic diseases such as excessive use of alcohol, cigarette smoking, obesity and unhealthy diets are unlikely to occur in an entirely uniform manner in a population. That is, it is uncertain whether people can be accurately classified into two groups, as those adapting to a healthy versus unhealthy lifestyle. For example, certain individuals who engage in more vigorous physical activities tend to consume alcohol more frequently than their inactive counterparts.¹⁵ Also, cigarette smoking is associated with lower rate of obesity among certain individuals.¹⁶ In a study among the US adults, Leventhal *et al* (2014) used modifiable risk factors like usage of alcohol, drug, nicotine, current obesity status and weekly physical activity and found that the adult population clusters into five subgroups based on the pattern of these modifiable risk factors.¹⁴ Another study in India by Shaikh and Khan (2021) found that hypertension is more likely to cluster with other modifiable risk behaviours such as smoking, alcohol, chewing tobacco, obesity and unhealthy diets in both men and women.¹⁷ Atorkey and Owiredua (2021), in their study that identified the clustering of five multiple health risk behaviours, also found that low vegetable and fruit intake and low physical activity cluster together within an individual, whereas smoking tobacco and alcohol consumption co-occur together forming another cluster.¹⁸ The pattern of clustering of modifiable risk behaviours could also differ by region and community level, as evident from previous studies.^{17 19} The clustering of two or more modifiable risk behaviours is of great concern because it can intensify the risk of developing chronic diseases and cardiovascular mortality,^{20 21 22} and most intervention measures are specific to a single risk factor. For example, in Asia 43.5% of Chinese adults and 37% of Korean adults had at least two cardiovascular risk factors.^{11 23} Moreover, in India, 35.6% of adult men and 10% adult women have at least three lifestyle risk factors.¹⁷ It is also found that a combination of smoking and heavy drinking emerged as the riskiest behaviour for all causes of death.²⁴

The majority of studies reported a higher prevalence of smoking, excess alcohol consumption, physical inactivity and poor diet among people with diabetes or hypertension than those without diabetes or hypertension.^{25 26}

Notably, previous studies have examined patients with diabetes or hypertension as a single homogeneous group without considering the possible heterogeneity existing within the population.²⁷ Most of the existing studies on health risk behaviours have emphasised only changing a single unhealthy behaviour.^{26 28} However, interventions that simultaneously target these unhealthy behaviours are necessary from the policy perspective.¹¹

Existing studies in India on clustering of NCD risk factors have used intraclass correlation and scoring methods.^{17 19 29} However, studies have suggested that latent class analysis (LCA) is the most appropriate method for explaining the clustering of health risk behaviours.^{30 31} LCA is a probabilistic model approach designed for identifying clusters based on dichotomous variables. This study is also an attempt to explore the clustering of NCD risk factors using the LCA approach in NE India. The aim is to identify the latent classes of modifiable risk factors among patients with diabetes and hypertension based on the observed indicator variables such as smoking, alcohol, aerated drinks, overweight or obesity, diabetes and hypertension. We applied LCA also to investigate heterogeneity in the population of an individual with diabetes and hypertension based on the responses of observed risk factors. More specifically, the objective is to describe how individuals in NE India diagnosed with diabetes or hypertension are clustered according to the pattern of observed risk factors and identify the difference between these groups based on sociodemographic characteristics. Information on clustering of health risk behaviours become vital as it guides in designing the disease control programmes. Intervention approach that aimed at targeting multiple modifiable risk behaviours in a single programme is gaining importance because they are potentially most cost-effective, efficient and may have a great public health impact than the single modifiable risk approaches. Therefore, we hypothesise that the study population diagnosed with diabetes or hypertension is homogeneous with respect to the modifiable risk factors.

METHODS

Study location

The study focuses on the NE region of India comprising eight states, namely Assam, Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim and Tripura (figure 1). Most parts of the region has a hilly terrain and inhabited mainly by myriad tribes belonging to different cultures and ethnic communities.³² The region has over 45 million populations, which is 3.8% of India's population, with a density of 159 persons/km².³³

Design and setting

The fourth round of the Indian National Family Health Survey (NFHS-4) (2015–2016), a nationally representative large-scale survey, is a cross-sectional study that adopted a stratified two-stage random sampling design for both urban and rural areas, including a systematic mapping

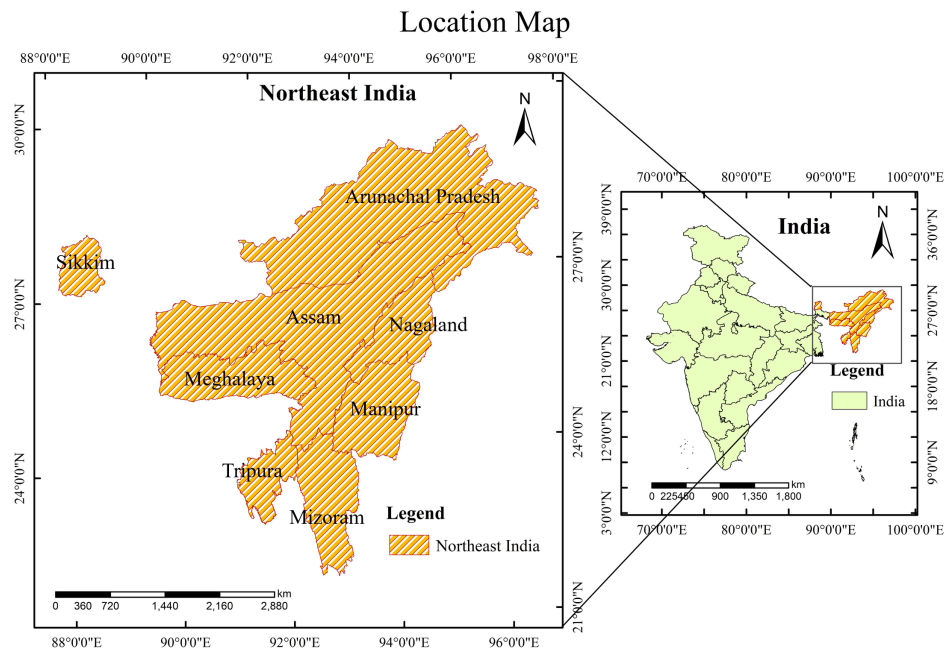


Figure 1 Map showing the location of the study area.

and listing of households in the selected primary sampling units (PSUs). The analysis is based on the publicly available data, in which the identity of selected respondents, households and geocodes of PSUs are concealed.³⁴ In the first stage, the PSUs were selected based on the probability proportional to population size. Rural PSUs are the villages, while census enumeration blocks formed the PSUs in urban areas. In the second stage, systematic random sampling was done in each PSUs to select households for the sample. At the household level, information were sought from the heads, eligible women aged 15–49 years and men aged 15–54 years, along with their biomarkers and blood samples on a specially designed dried blood spots. A detailed description of the survey design can be found elsewhere.³⁴

The data used in this study is the NFHS-4 conducted during 2015–2016. The data were downloaded from the Demographic Health Survey (DHS) website.³⁵ The NFHS series of surveys has been initiated by the Ministry of Health and Family Welfare, Government of India, and coordinated by the International Institute for Population Sciences, Mumbai.

Sample

The study includes only individuals aged 15–49 years who were diagnosed with either diabetes or hypertension or both. The total sample size is 22 249, of which 3284 were men and 18 965 were women from the NE region of India.

Patient and public involvement

No patient involved.

Measures

Diabetes

Fingerstick blood is collected, and blood glucose level is determined using the FreeStyle Optium H Glucometer. A

respondent is considered diagnosed with diabetes if the random blood sugar level is >140 mg/dL.³⁴

Hypertension

BP level was measured using an OMRON BP measuring device. Three separate BP readings were taken with an interval of 5 min between readings. A respondent is considered hypertensive if the systolic BP is ≥ 140 mm Hg or diastolic BP ≥ 90 mm Hg or if the respondent was taking antihypertensive medication to lower BP at the time of the survey.³⁴

Overweight or obesity

In this study, we categorised continuous body mass index (BMI) according to the WHO guidelines for the Asian population. A respondent is classified as overweight or obese if BMI is ≥ 23 kg/m².³⁶

Outcome variables of interest

Observed indicator variables

The observed variables used as latent indicators are cigarette smoking (no=1, yes=2), alcohol consumption (no=1, yes=2), taking aerated drinks (no=1, yes=2), overweight or obesity (no=1, yes=2), diabetes (no=1, yes=2) and hypertension (no=1, yes=2).

Predictors or concomitant variables

Concomitant variables for class determination

The concomitant variables include age, gender, education, marital status and household wealth index. Figure 2 illustrates the conceptual framework of the study.

Statistical analysis

LCA, based on structural equation modelling, is a person-centred analytical approach that allows the identification of homogeneous subgroups in a heterogeneous

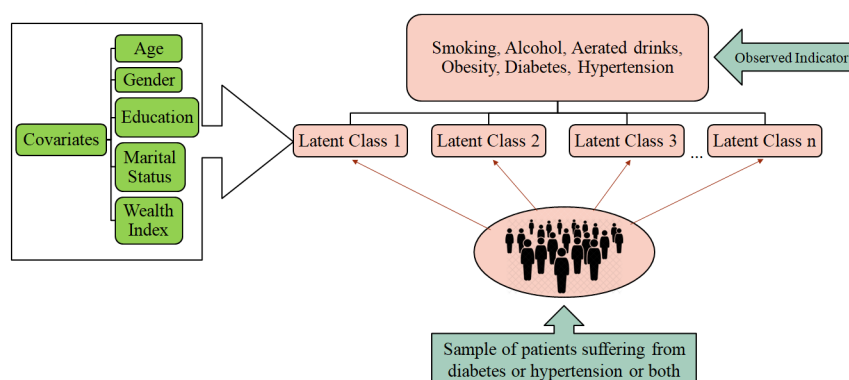


Figure 2 Conceptual framework of the study.

population. LCA is similar to cluster analysis because individuals are classifying into homogeneous unobserved (latent) groups based on the response pattern to a set of observed variables. However, in cluster analysis, objects are classified based on distance measures, whereas in LCA classification are based on probabilistic and finite mixture modelling approach.^{37 38} A latent class model splits the population into mutually exclusive and exhaustive groups that are homogeneous within the groups but differ among them. Several studies have adopted the LCA approach to identify risk factors clustering within the heterogeneous population. In a study by Dey *et al* (2016), based on seven observed risk factors, LCA was used to classify individuals into two latent classes ‘susceptible to adverse health outcomes groups’ and ‘not susceptible to adverse health outcomes groups’.³⁹ Another study in China applied LCA to 10 observed complications and comorbidities of patients with type 2 diabetes and categorised the individuals into four latent classes, namely ‘complications and comorbidity groups’, ‘high risk of complications group’, ‘high risk of comorbidities and Cardiovascular Disease groups’ and ‘diabetes without complications and comorbidities group’.⁴⁰ A study in West Azerbaijan province among the patients with hypertension aged 50 years and above used four indicators, such as dietary patterns, physical activity, tobacco use and high BP control to categorise the patients with hypertension into three latent classes.⁴¹

LCA is an innovative statistical method used to identify the latent classes of homogeneous individuals in a heterogeneous population.⁴⁰ The latent class was performed using the six indicator variables mentioned above. LCA uses observed dichotomous indicators to identify the unobserved latent class in a heterogeneous population. To determine the optimal number of latent classes, we examined one to four models. The optimal number of latent classes was chosen based on the model having the smallest Akaike Information Criterion (AIC), consistent Akaike Information Criterion (cAIC), Bayesian-Schwarz Information Criterion (BIC), adjusted Bayesian-Schwarz Information Criterion and entropy.^{42 43} A combination of parsimony and interpretability model selection criteria was used to carefully select the number of latent classes

so that individuals are allocated to their most likely class. The outputs of the LCA include the number of latent classes, the latent class probability (ie, the probability that an individual selected at random belonged to each latent class) and the conditional probability (ie, the probability that an individual would give a particular response to a specific item of an observed indicator variable given that an individual belongs to a specific latent class).

To determine the association between the concomitant variables (ie, age, gender, education, marital status and household wealth index) and the latent class, we apply a one-step technique approach. In this approach, the concomitant variables are included in the latent class regression model, and their coefficient is estimated simultaneously as part of the latent class model.⁴⁴ This approach has been demonstrated to provide the best and unbiased coefficient estimates of the concomitant variables as compared with other methods.⁴⁵ Statistical analysis was conducted in R statistical software, and LCA was analysed using the poLCA package (Polytomous Variable Latent Class Analysis).⁴⁴

RESULTS

Descriptive statistics

A total of 22 249 individuals with chronic diseases, that is, diagnosed with diabetes or hypertension, were included in the analysis (table 1). Individuals in the age group of 45–49 years and 15–19 years constitute the largest (21.3%) and the lowest (5.0%), respectively, in the total sample. Over half (51.2%) of the sample have completed secondary education, 24.6% were illiterate and 79.6% were currently married. Further, among the patients with chronic diseases, the prevalence of smoking, consumption of alcohol and taking aerated drinks was 4.2%, 14.8% and 68.6%, respectively. Also, 81.1% of the participants reported suffering from hypertension and 27.7% from diabetes. The sample distribution also indicates that chronic condition is much higher among women (86.1%) than men (13.9%), among older ages (35+ years) and among poorer households.

Table 1 Per cent distribution of adults aged 15–49 years with chronic condition (patients with diabetes and hypertension) by selected background characteristics, Northeastern India, NFHS-4, 2015–2016

Background characteristics	% (n=22249)
Age, years	
15–19	5.0
20–24	8.4
25–29	12.2
30–34	15.0
35–39	19.2
40–44	19.0
45–49	21.3
Sex	
Male	13.9
Female	86.1
Education	
Illiterate	24.6
Primary	16.3
Secondary	51.2
Higher secondary	8.0
Marital status	
Never married	13.6
Married	79.6
Widowed/divorced/separated	6.87
Wealth index	
Poorest	16.7
Poorer	34.6
Middle	22.4
Richer	17.5
Richest	8.8
Observed indicators	
Smoking	4.2
Drink alcohol	14.8
Take aerated drinks	68.6
Overweight or obesity	43.3
Diabetes	27.7
Hypertension	81.1

NFHS-4, fourth round of the Indian National Family Health Survey.

Model fit and selection of latent class

The model fit/selection statistics to derive appropriate models are shown in table 2. We fit the LCA models with classes ranging from 1 to 4. For each LCA model, the model fit indicators BIC, AIC, cAIC and entropy were calculated. According to the model fit indicators, a three-class model has been selected as the best fit model as it has the lowest BIC and cAIC values. Smaller values on each indicator suggest a better model, or a model with

few explanatory variables or parameters, and inform the decision of the best model to be retained.⁴⁶ Another diagnostic indicator is entropy⁴⁷ which indicates how accurately the model defines the classes. In general, an entropy value close to 1.0 is considered as ideal⁴⁸ and above 0.8 is also acceptable. An entropy value of 1.0 would indicate that every individual has been perfectly classified within the classes based on the responses of the observed indicators. The three latent classes were labelled as ‘diabetic with low-risk lifestyle’, ‘high-risk lifestyle’ and ‘hypertensive with low-risk lifestyle’. Figure 3 portrays the graphical representation of item response probability for the indicator’s variables across the latent classes.

Latent class probability and the conditional probability of a ‘yes’ response for each indicator variables were summarised in table 3. The last row in table 3 indicates the probability of class membership in each latent class. About 21% of the participants were expected to belong to the class 1 ‘diabetic with low-risk lifestyle’, 8% to class 2 ‘high-risk lifestyle’ and 71% to class 3 ‘hypertensive with low-risk lifestyle’. Members in class 1 were likely to report having high blood sugar level (diabetic) and unlikely to smoke cigarette or drinking alcohol. Members in class 2 are more likely to smoke cigarette and drink alcohol, and members in class 3 are likely to report having high BP and unlikely to smoke cigarette or drink alcohol.

Table 4 presents the distribution of socioeconomic and demographic characteristics of the individuals by the three latent classes. Age, gender, education, marital status and household wealth index were statistically found significantly associated with the three latent classes at 1% level of significant. The hypertensive with low-risk lifestyle class has the highest proportion of respondents aged 45–49 years, whereas diabetic with both low-risk lifestyle class and high-risk lifestyle class has the highest proportion of respondents aged 35–39 years (19.4% and 19.5%, respectively). Both the diabetic with low-risk lifestyle class (10.6% men vs 89.4% women) and the hypertensive with low-risk lifestyle class (6.2% men vs 93.8% women) has the highest proportion of women; men constituted almost 100% of the high-risk lifestyle class (99.2% men vs 0.8% women). Lower proportion of the respondents have completed higher secondary education in case of both diabetic with low-risk lifestyle class (10.5%) and hypertensive with low-risk lifestyle class (8.5%). Marital status shows a similar pattern in all the three classes; higher proportion of married respondents and lower proportion of widowed/divorced/separated respondents (<10%). The wealth index shows that respondents belonging to the middle quintile constituted the largest proportion in the diabetic with low-risk lifestyle class (27.2%) and high-risk lifestyle class (26.6%). However, respondents belonging to the poorer wealth quintile have the highest proportion of hypertensive with low-risk lifestyle class (29.4%).

Covariates predicting latent class membership

Table 5 summarised the OR results from the latent regression model. Compared with respondents in the 15–19

Table 2 Model fit statistics for the latent class models (n=22 249)

No of classes	Log likelihood	Residual df	BIC	aBIC	cAIC	Likelihood ratio	Entropy
1	-69173.1	57	138406.3	138387.2	138412.3	17350.8	-
2	-61315.4	34	122921.0	122828.9	122950.0	1859.2	1.0
3	-58920.2	11	118361.0	118195.7	118413.0	1168.6	0.9
4	-59970.1	-12	120691.0	120452.6	120766.0	2396.1	0.9

aBIC, adjusted Bayesian-Schwarz Information Criterion; BIC, Bayesian-Schwarz Information Criterion; cAIC, consistent Akaike Information Criterion; df, degrees of freedom.

years age group, those in the 20–24 age group were 4.1 times (OR=4.10; $p<0.01$) more likely to belong to class 2 (high-risk lifestyle) than to class 1 (diabetic with low-risk lifestyle). Similarly, compared with respondents in the 15–19 year age group, those in the 45–49 age group were 1.7 times (OR=1.73; $p<0.01$) more likely to belong to class 3 (hypertensive with low-risk lifestyle) than to class 1 (diabetic with low-risk lifestyle). Also, we observed that the odds of being in class 3 increases with the age of the respondents. With respect to men, women were less likely (OR=0.002; $p<0.01$) to belong to class 2 but more likely (OR=1.46; $p<0.01$) to belong to class 3. Higher odds of being in class 2–3 than to class 1 were observed among the married respondents (OR=1.54 and OR=1.16). Similarly, widowed/divorced/separated were more likely to be in class 2 (OR=3.35; $p<0.01$), but less likely to be in class 3 (OR=0.98; $p>0.05$). Moreover, compared with the illiterate respondents, those with higher secondary education were less likely to belong to class 3 than to class 1 (OR=0.79; $p<0.01$), and being in the richest wealth quintile lower is the odds of being in class 3 than in class 1 (OR=0.69; $p<0.01$).

DISCUSSION

This study is the first attempt in NE India, a region characterised by a seemingly homogeneous racial group but inhabited by ethnolinguistically diverse subpopulations that applied LCA to provide evidence about the pattern of modifiable risk factors among people suffering from diabetes or hypertension. Based on the observed modifiable health risk behaviours, the LCA suggests presence of three classes of respondents with diabetes or hypertension: ‘diabetic with low-risk lifestyle’, ‘high-risk lifestyle’, and ‘hypertensive with low-risk lifestyle’. The findings portray that the 22 249 odd diabetic or hypertensive persons in the region can be categorised into three groups or classes, with homogeneity in their characteristics within the group but heterogeneous between the groups. One of the three latent classes has a 100% likelihood of having one of the risk factors combined with a low and moderate probability of having the other five risk factors. Each of the classes exhibit a unique risk factor configuration, and socioeconomic and demographic profiles from one another. Thus, chronic disease prevention programmes that specially target and reach out to

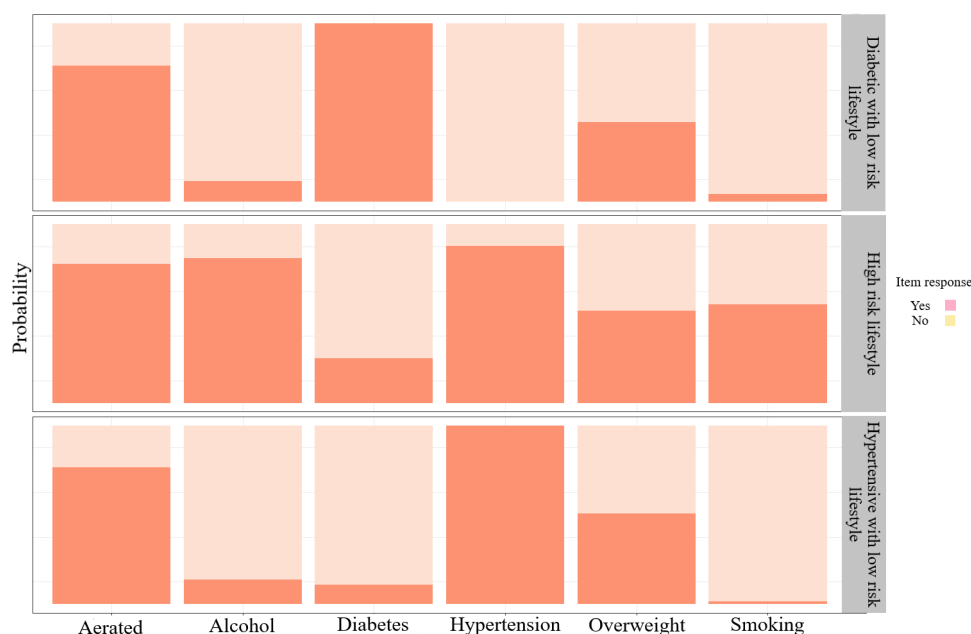
**Figure 3** Item response conditional probability across the latent classes.

Table 3 The conditional probability of the response items and the latent class probability for study sample in Northeastern India, NFHS-4, 2015–2016

Indicator variables	Diabetic with low-risk lifestyle (class 1)	High-risk lifestyle (class 2)	Hypertensive with low-risk lifestyle (class 3)
Smoking	0.04	0.55	0.02
Drink alcohol	0.12	0.81	0.14
Take aerated drinks	0.76	0.78	0.77
Overweight or obesity	0.44	0.52	0.51
Diabetic	1.00	0.25	0.11
Hypertensive	0.00	0.88	1.00
Latent class probability	0.21	0.08	0.71

NFHS-4, fourth round of the Indian National Family Health Survey.

the at-risk subpopulations will be beneficial and effective, as illustrated in this analysis (tables 3 and 5).

More than 70% of the respondents belong to the class ‘hypertensive with low-risk lifestyle’ have their BP level higher than normal and less likely to smoke and use alcohol. Evidence from a large body of existing literature supports that cigarette smoking leads to an immediate increase in BP level and heart rate.⁴⁹ The effect of cigarette smoking on BP and heart rate is largely due to the nicotine content in the cigarettes.⁴⁹ However, studies also found a significant elevation in BP and an increase in the prevalence of hypertension following smoking cessation,^{50 51} which could possibly be due to postcessation weight gain.^{49 50} Green *et al*⁵² had concluded that smoking cessation in itself is not likely to result in lowering BP but on the contrary it may result even in higher BP.⁵² Moreover, about 8% of the respondents belonging to the ‘high-risk lifestyle’ class mostly drink alcohol and smoke cigarettes. This co-occurrence of alcohol use and cigarette smoking has been well documented in the existing studies,^{53 54} and evidence suggest that a common genetic trait play a role in the co-occurrence of smoking and alcohol use.^{55 56} Epidemiological studies have also shown that individuals who are dependent on nicotine are more than four time more likely to be dependent on alcohol. Similarly, individuals who are dependent on alcohol are also more than three times more likely to be dependent on nicotine with reference to the general population.⁵⁷ Stress is another factor that influences use of tobacco and alcohol. For example, people who drink alcohol are most likely to smoke in an attempt to overcome stress.⁵⁸ Literature also suggest cross-tolerance effect between alcohol and tobacco. It is found that the stimulating effect of nicotine in tobacco is restrained by the sedating effect of alcohol which entices smokers to drink more alcohol.⁵⁹ Alcohol’s sedating effect reduces the physiological effect of nicotine, thus increasing the consumption of tobacco.

In all the three identified classes, a vast majority of diabetic or hypertensive persons appear to have a habit of consuming aerated drinks and an uncontrolled body weight. Studies suggest that sugar-sweetened beverage intake is the major contributor to weight gain and can increase the risk of T2 diabetes and hypertension.^{60 61}

Cross-sectional studies acknowledge the positive association between overweight and obesity with the consumption of aerated drinks or sugar-sweetened beverages. For example, studies infer that women who consumed sweetened beverages more than once a week were more likely to be overweight than those who consumed sweetened beverages less than once a week.⁶² Other studies found that women who consumed sugar-sweetened beverages regularly were heavier by 0.2 kg than the non-consumers.⁶³ The intake of sodas sweetened with high-fructose corn syrup is associated with hypertension among adolescent population.⁶⁴ Also, a spurt in the prevalence of hypertension has been observed among individuals who consumed more than average amount of fructose.⁶⁵ Evidence from a cross-sectional study in NE India supports the positive correlation between overweight or obesity with diabetes and hypertension.^{66 27} The plausible explanation is that both sugar sweetened beverages and artificial sweetened beverages are associated with the development of metabolic derangements that in turn might lead to elevated BP.⁶⁷ High intake of sugar-sweetened beverages contributes to a higher total energy consumption which increases energy intake without an increase in energy expenditure certainly results in weight gain. There is no dearth of scientific evidence which support the notion that decreasing aerated drinks consumption reduces the prevalence of overweight and obesity.

As observed in a previous study,⁶⁸ this study also found that those aged 20 years and above are more likely to be in the ‘high-risk cluster’ than those below 20 years. Studies have elucidated that the adolescence or teenage stage is just the initial stage when an individual starts using or exposed to alcohol or tobacco (smoking) products. However, as this behaviour becomes a habit later on, it subsequently accentuates the intensity of substance abuse among older adults.^{69 70} Also, it is well documented that initiation of smoking and drinking alcohol at an early age is associated with the number of cigarettes smoked and the quantity of alcohol consumed per day in adult age. According to Investigator (1991) and Dawson *et al* (2008), individuals who began to smoke or drink before the age of 20 years were most likely to consume a large volume of alcohol and smoke more cigarettes per day as compared with those who started after the age of 20 or above.^{71 72} Further, compared with young adults, older adults were more likely to report high BP and less likely to smoke or drink alcohol. One reason could be the existing condition of high BP; an older adult was most likely to have taken a precaution by lowering or quitting smoking and alcohol consumption to prevent further deterioration of health. Evidence from a recent study among older adults demonstrated that the main reason for smoking cessation attempt

Table 4 Per cent distribution of patients with diabetes and hypertension aged 15–49 years according to three latent classes by background characteristics, Northeastern India, NFHS-4, 2015–2016

Background characteristics	'Diabetic with low-risk lifestyle' class (n=4727, 21%)	'High-risk lifestyle' class (n=1828, 8%)	'Hypertensive with low-risk lifestyle' class (n=15 694, 71%)	P value*
Age, years				<0.01
15–19	8.08	3.06	4.96	
20–24	8.99	8.37	7.6	
25–29	12.88	13.51	11.93	
30–34	15.30	18.00	15.29	
35–39	19.40	19.53	19.46	
40–44	17.73	19.09	19.04	
45–49	17.62	18.44	21.72	
Gender				<0.01
Male	10.62	99.18	6.17	
Female	89.38	0.82	93.83	
Education				<0.01
Illiterate	17.96	11.65	25.17	
Primary	16.50	15.10	15.85	
Secondary	55.05	58.70	50.50	
Higher secondary	10.49	14.55	8.47	
Marital status				<0.01
Never married	19.93	21.55	14.30	
Married	72.75	75.22	78.41	
Widowed/divorced/separated	7.32	3.23	7.30	
Wealth index				<0.01
Poorest	10.94	12.14	12.94	
Poorer	26.06	24.84	29.43	
Middle	27.23	26.64	26.30	
Richer	22.49	23.34	21.25	
Richest	13.29	12.04	10.07	
Total	100.0	100.0	100.0	

*Based on χ^2 tests of association.

NFHS-4, fourth round of the Indian National Family Health Survey.

was the motivation towards better health.⁷³ Hypertensive smokers were encouraged to quit smoking because of the risk of developing a severe form of hypertension such as malignant and renovascular hypertension.⁷⁴ Gender factors are also found to be significantly associated with class 2 'high-risk lifestyle' and class 3 'hypertensive with low-risk lifestyle'. Consistent with the previous studies, in this study also, female respondents were less likely to be in the high-risk cluster than male respondents.⁷⁵ One primary reason stated is that smoking and drinking alcohol by women is not socially and culturally accepted by many,⁷⁶ and women are more concerned about their health and tend to avoid unhealthy lifestyle behaviours.⁷⁷ In our study, individuals from a wealthier household with higher education were less likely to belong to a cluster of hypertensive with low substance use. An individual with

higher education and better income can afford to pay for a healthier lifestyle, including regular physical exercise, accessibility to advanced and quality healthcare services, as such efforts may reduce the risk of hypertension.^{66 78 79} Being widowed/divorced/separated tend to increase the likelihood of falling into the high-risk cluster. Substance use by marital status indicated that widowed or divorced individuals are most likely to consume alcohol, cigarette and marijuana as compared with singles.^{80 81}

A search of similar studies indicates that no study in India or NE India has examined the clustering of health risk factors using the latent class approach. However, our study has some limitations. First, the cross-sectional design of the study restricts us to measure the causal effect. Second, due to the dichotomous response on alcohol consumption, we cannot measure the quantity of

Table 5 The odds for various latent classes of lifestyle behaviours and their associated covariates among individuals aged 15–49 years, Northeastern India, NFHS-4, 2015–2016

Background characteristics	Class 2 vs class 1		Class 3 vs class 1	
	OR	SE	OR	SE
Age, years				
15–19	Ref		Ref	
20–24	4.10***	0.25	1.27**	0.09
25–29	5.10***	0.25	1.34***	0.09
30–34	5.70***	0.26	1.42***	0.09
35–39	3.53***	0.26	1.42***	0.09
40–44	3.82***	0.27	1.51***	0.09
45–49	3.46***	0.27	1.73***	0.09
Gender				
Male	Ref		Ref	
Female	0.002***	0.34	1.46***	0.07
Education				
Illiterate	Ref		Ref	
Primary	1.03	0.19	0.73***	0.06
Secondary	1.08	0.17	0.79***	0.05
Higher secondary	0.70	0.21	0.79***	0.08
Marital status				
Never married	Ref		Ref	
Married	1.54**	0.15	1.16**	0.06
Widowed/ divorced/ separated	3.35***	0.35	0.98	0.08
Wealth index				
Poorest	Ref		Ref	
Poorer	0.90	0.18	1.00	0.06
Middle	1.02	0.18	0.87**	0.07
Richer	1.14	0.19	0.86*	0.07
Richest	1.13	0.22	0.69***	0.08

***p<0.01; **p<0.05; *p<0.1.
Ref, reference group (class 1).

alcohol an individual consumed. Lastly, information on social groups and occupations could not be included in the analysis because of the large number of missing cases in these variables.

CONCLUSION

In NE India, as this study illustrates, there are three classes of modifiable risk factors closely related to the risk of chronic diseases (diabetes and hypertension). In addition, the study also identified factors that uniquely distinguished the identified classes. The intake of aerated drinks and obesity are the common modifiable risk factors in all three classes. As an initiative to healthy lifestyle behaviours, the findings suggest that men should control smoking and alcohol consumption, while women should

control body weight and BP. Further, we found that marital status and gender could be the catalyst to prevent high-risk smoking and alcohol drinking in NE India. Also, the study observes smoking and alcohol use are the two modifiable risk behaviours, which tends to co-occur within an individual. Therefore, it is recommended that policy and intervention programmes in NE India that promote healthy lifestyles should focus on targeting multiple modifiable risk behaviours that are most likely to co-occur within an individual. On the other hand, awareness about adapting to healthy diets and weight control along with physical exercises should be promoted with rigour in the region.

Correction notice This article has been corrected since it was first published. Author affiliation section has been updated.

Acknowledgements The authors are thankful to the editor and the two reviewers for their insightful comments and valuable suggestions which hugely improved the analysis and the overall quality of the paper.

Contributors Conceived and designed the research paper: SPM, HL. Analysed the data: SPM, HSC. Contributed agents/materials/analysis tools, drafting manuscript and refined the manuscript: all authors. guarantor: SPM

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval This study does not involve human participants.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Data stored and maintained by the DHS office and available to public at https://dhsprogram.com/data/dataset/India_Standard-DHS_2015.cfm?flag=1 upon request.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Strong P Marbaniang <http://orcid.org/0000-0001-5347-1867>

REFERENCES

- 1 Chobanian AV, Bakris GL, Black HR, *et al*. The seventh report of the joint National Committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. *JAMA* 2003;289:2560–72.
- 2 International Diabetes Federation. *Advocacy guide to the IDF diabetes atlas*. 9th edition. Brussels, Belgium: International Diabetes Federation, 2019. <http://www.diabetesatlas.org/>
- 3 Ford ES, Zhao G, Tsai J, *et al*. Low-Risk lifestyle behaviors and all-cause mortality: findings from the National health and nutrition examination survey III mortality study. *Am J Public Health* 2011;101:1922–9.
- 4 Geldsetzer P, Manne-Goehler J, Theilmann M, *et al*. Geographic and sociodemographic variation of cardiovascular disease risk in India: a cross-sectional study of 797,540 adults. *PLoS Med* 2018;15:e1002581.
- 5 Ghosh S, Kumar M. Prevalence and associated risk factors of hypertension among persons aged 15–49 in India: a cross-sectional study. *BMJ Open* 2019;9:e029714.
- 6 Gupta R, Gaur K, S Ram CV. Emerging trends in hypertension epidemiology in India. *J Hum Hypertens* 2019;33:575–87.

- 7 India State-Level Disease Burden Initiative Diabetes Collaborators. The increasing burden of diabetes and variations among the states of India: the Global Burden of Disease study 1990-2016. *Lancet Glob Health* 2018;6:e1352-62.
- 8 WHO. *Global status report on noncommunicable diseases 2014*. World Health Organization, 2014.
- 9 Kvaavik E, Batty GD, Ursin G, et al. Influence of individual and combined health behaviors on total and cause-specific mortality in men and women: the United Kingdom health and lifestyle survey. *Arch Intern Med* 2010;170:711.
- 10 Loef M, Walach H. The combined effects of healthy lifestyle behaviors on all cause mortality: a systematic review and meta-analysis. *Prev Med* 2012;55:163-70.
- 11 Ha S, Choi HR, Lee YH. Clustering of four major lifestyle risk factors among Korean adults with metabolic syndrome. *PLoS One* 2017;12:e0174567.
- 12 Poortinga W. The prevalence and clustering of four major lifestyle risk factors in an English adult population. *Prev Med* 2007;44:124-8.
- 13 Schuit AJ, van Loon AJM, Tijhuis M, et al. Clustering of lifestyle risk factors in a general adult population. *Prev Med* 2002;35:219-24.
- 14 Leventhal AM, Huh J, Dunton GF. Clustering of modifiable biobehavioral risk factors for chronic disease in US adults: a latent class analysis. *Perspect Public Health* 2014;134:331-8.
- 15 Lisha NE, Martens M, Leventhal AM. Age and gender as moderators of the relationship between physical activity and alcohol use. *Addict Behav* 2011;36:933-6.
- 16 Akbartabtoori M, Lean MEJ, Hankey CR. Relationships between cigarette smoking, body size and body shape. *Int J Obes* 2005;29:236-43.
- 17 Shaikh R, Khan J. Clustering of lifestyle risk factors among adult population in India: a cross-sectional analysis from 2005 to 2016. *PLoS One* 2021;16:e0244559.
- 18 Atorkey P, Owiredua C. Clustering of multiple health risk behaviours and association with socio-demographic characteristics and psychological distress among adolescents in Ghana: a latent class analysis. *SSM Popul Health* 2021;13:100707.
- 19 Bischops AC, De Neve J-W, Awasthi A, et al. A cross-sectional study of cardiovascular disease risk clustering at different socio-geographic levels in India. *Nat Commun* 2020;11:5891.
- 20 Marbaniang SP, Lhungdim H, Chauhan S, et al. Interaction of multiple risk factors and population attributable fraction for type 2 diabetes and hypertension among adults aged 15-49 years in northeast India. *Diabetes Metab Syndr* 2021;15:102227.
- 21 Laxer RE, Brownson RC, Dubin JA, et al. Clustering of risk-related modifiable behaviours and their association with overweight and obesity among a large sample of youth in the COMPASS study. *BMC Public Health* 2017;17:102.
- 22 Lee DH, Nam JY, Kwon S, et al. Lifestyle risk score and mortality in Korean adults: a population-based cohort study. *Sci Rep* 2020;10:10260.
- 23 Gao B, Zhang L, Wang H, et al. Clustering of major cardiovascular risk factors and the association with unhealthy lifestyles in the Chinese adult population. *PLoS One* 2013;8:e66780.
- 24 Hart CL, Davey Smith G, Gruer L, et al. The combined effect of smoking tobacco and drinking alcohol on cause-specific mortality: a 30 year cohort study. *BMC Public Health* 2010;10:789.
- 25 Sobngwi E, Mbanya J-CN, Unwin NC, et al. Physical activity and its relationship with obesity, hypertension and diabetes in urban and rural Cameroon. *Int J Obes* 2002;26:1009-16.
- 26 Aeri BT. Risk factors associated with the increasing cardiovascular diseases prevalence in India: a review. *J Nutr Food Sci* 2014;5:1-6.
- 27 Marbaniang SP, Lhungdim H, Yadav B, et al. Overweight/obesity risks and prevalence of diabetes and hypertension in North eastern India: an analysis using seemingly unrelated probit model. *Clin Epidemiol Glob Health* 2021;11:100764.
- 28 Umamahesh K, Vigneswari A, Surya Thejaswi G, et al. Incidence of cardiovascular diseases and associated risk factors among subjects with type 2 diabetes - an 11-year follow up study. *Indian Heart J* 2014;66:5-10.
- 29 Sarveswaran G, Kulothungan V, Mathur P. Clustering of noncommunicable disease risk factors among adults (18-69 years) in rural population, South-India. *Diabetes Metab Syndr* 2020;14:1005-14.
- 30 Schreiber JB. Latent class analysis: an example for reporting results. *Res Social Adm Pharm* 2017;13:1196-201.
- 31 Weller BE, Bowen NK, Faubert SJ. Latent class analysis: a guide to best practice. *J Black Psychol* 2020;46:287-311.
- 32 Marbaniang SP. Women care and practices in the management of childhood diarrhea in northeast India. *Child Care Pract* 2020;3:1-13.
- 33 Dikshit KR, Dikshit JK. Population of the North-Eastern States of India. In: *North-East India: land, people and economy. advances in Asian Human-Environmental research*. Springer: Dordrecht;, 2014: 421-56.
- 34 International Institute for Population Sciences (IIPS) and ICF. *National family health survey (NFHS-4), 2015-16: India*. Mumbai: IIPS, 2017. <http://rchiips.org/nfhs/NFHS-4Reports/India.pdf>
- 35 DHS. The DHS Program-India: standard DHS, 2015-16 dataset,. Demogr. heal. Surv, 2017. Available: https://dhsprogram.com/data/dataset/India_Standard-DHS_2015.cfm?flag=1
- 36 WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* 2004;363:157-63.
- 37 Islam MM, Valderas JM, Yen L, et al. Multimorbidity and comorbidity of chronic diseases among the senior Australians: prevalence and patterns. *PLoS One* 2014;9:e83783.
- 38 Brusco MJ, Shireman E, Steinley D. A comparison of latent class, k-means, and K-median methods for clustering dichotomous data. *Psychol Methods* 2017;22:563-80.
- 39 Dey A, Chakraborty A, Majumdar K, et al. Application of latent class analysis to estimate susceptibility to adverse health outcomes based on several risk factors. *Int J Community Med Public Health* 2016;3:3423-9.
- 40 Gao F, Chen J, Liu X, et al. Latent class analysis suggests four classes of persons with type 2 diabetes mellitus based on complications and comorbidities in Tianjin, China: a cross-sectional analysis. *Endocr J* 2017;64:1007-16.
- 41 Ghanbari J, Mohammadpoorasl A, Jahangiry L, et al. Subgroups of lifestyle patterns among hypertension patients: a latent-class analysis. *BMC Med Res Methodol* 2018;18:127.
- 42 Sugiura N. Further analysts of the data by akaike' s information criterion and the finite corrections. *Commun Stat Theory Methods* 1978;7:13-26.
- 43 Akaike H. A new look at the statistical model identification 1974:215-22.
- 44 Linzer DA, Lewis JB. polCA : An R Package for Polytomous Variable Latent Class Analysis. *J Stat Softw* 2011;42.
- 45 Bolck A, Croon M, Hagenaaers J. Estimating latent structure models with categorical variables: one-step versus three-step estimators. *Polit. Anal* 2004;12:3-27.
- 46 Scotto Rosato N, Baer JC. Latent class analysis: a method for capturing heterogeneity. *Soc Work Res* 2012;36:61-9.
- 47 Wang M-C, Deng Q, Bi X, et al. Performance of the entropy as an index of classification accuracy in latent profile analysis: a Monte Carlo simulation study. *Acta Psychologica Sinica* 2017;49:1473.
- 48 Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. *J Classif* 1996;13:195-212.
- 49 Omvik P. How smoking affects blood pressure. *Blood Press* 1996;5:71-7.
- 50 Jenner DA, Puddey IB, Beilin LJ. Lifestyle- and occupation-related changes in blood pressure over a six-year period in a cohort of working men. *J Hypertens Suppl Off J Int Soc Hypertens* 1988;6:S605-7.
- 51 Gerace TA, Hollis J, Ockene JK, et al. Smoking cessation and change in diastolic blood pressure, body weight, and plasma lipids. MRFIT Research Group. *Prev Med* 1991;20:602-20.
- 52 Green MS, Jucha E, Luz Y. Blood pressure in smokers and nonsmokers: epidemiologic findings. *Am Heart J* 1986;111:932-40.
- 53 Ho RKS, Fok PWY, Chan HCH. Pattern and determinants of alcohol and tobacco co-use and its relationship with smoking cessation in Hong Kong. *Tob Prev Cessat* 2021;7:1-9.
- 54 Jiang N, Lee YO, Ling PM. Association between tobacco and alcohol use among young adult bar patrons: a cross-sectional study in three cities. *BMC Public Health* 2014;14:500.
- 55 Slutske WS, Eisen S, True WR, et al. Common genetic vulnerability for pathological gambling and alcohol dependence in men. *Arch Gen Psychiatry* 2000;57:666-73.
- 56 Bierut LJ, Schuckit MA, Hesselbrock V, et al. Co-Occurring risk factors for alcohol dependence and habitual smoking. *Alcohol Res Health* 2000;24:233-41.
- 57 Grant BF, Hasin DS, Chou SP, et al. Nicotine dependence and psychiatric disorders in the United States. *Arch Gen Psychiatry* 2004;61:1107.
- 58 Lyon RJ, Tong JE, Leigh G, et al. The influence of alcohol and tobacco on the components of choice reaction time. *J Stud Alcohol* 1975;36:587-96.
- 59 de Fiebre CM, Collins AC. A comparison of the development of tolerance to ethanol and cross-tolerance to nicotine after chronic ethanol treatment in long- and short-sleep mice. *J Pharmacol Exp Ther* 1993;266:1398-406.
- 60 Malik VS, Popkin BM, Bray GA, et al. Sugar-Sweetened beverages, obesity, type 2 diabetes mellitus, and cardiovascular disease risk. *Circulation* 2010;121:1356-64.

- 61 Qin P, Li Q, Zhao Y, *et al.* Sugar and artificially sweetened beverages and risk of obesity, type 2 diabetes mellitus, hypertension, and all-cause mortality: a dose-response meta-analysis of prospective cohort studies. *Eur J Epidemiol* 2020;35:655–71.
- 62 Liebman M, Pelican S, Moore SA, *et al.* Dietary intake, eating behavior, and physical activity-related determinants of high body mass index in rural communities in Wyoming, montana, and Idaho. *Int J Obes Relat Metab Disord* 2003;27:684–92.
- 63 French SA, Jeffery RW, Forster JL, *et al.* Predictors of weight change over two years among a population of working adults: the healthy worker project. *Int J Obes Relat Metab Disord* 1994;18:145–54.
- 64 Nguyen S, Choi HK, Lustig RH, *et al.* Sugar-Sweetened beverages, serum uric acid, and blood pressure in adolescents. *J Pediatr* 2009;154:807–13.
- 65 Jalal DI, Smits G, Johnson RJ, *et al.* Increased fructose associates with elevated blood pressure. *J Am Soc Nephrol* 2010;21:1543–9.
- 66 Marbaniang SP, Chungkham HS, Lungdim H. A structured additive modeling of diabetes and hypertension in northeast India. *PLoS One* 2022;17:e0262560.
- 67 Cohen L, Curhan G, Forman J. Association of sweetened beverage intake with incident hypertension. *J Gen Intern Med* 2012;27:1127–34.
- 68 Saikia N, Debbarma B. The socioeconomic correlates of substance use among male adults in northeast India. *Clin Epidemiol Glob Health* 2020;8:149–57.
- 69 IOM (Institute of Medicine). Patterns of Tobacco Use by Adolescents and Young Adults. In: Richard J, Kathleen Stratton LYK, eds. *Public health implications of raising the minimum age of legal access to tobacco products*. Washington, D.C: The National Academies Press, 2015. https://www.ncbi.nlm.nih.gov/books/NBK310412/pdf/Bookshelf_NBK310412.pdf
- 70 Maggs JL, Schulenberg JE. *Initiation and Course of Alcohol Consumption among Adolescents and Young Adults. In: Recent Developments in Alcoholism*. New York: Kluwer Academic Publishers-Plenum Publishers, 2005: 29–47.
- 71 Taioli E, Wynder EL, Investigators S. Effect of the age at which smoking begins on frequency of smoking in adulthood. *N Engl J Med* 1991;325:968–9.
- 72 Dawson DA, Goldstein RB, Chou SP, Patricia Chou S, *et al.* Age at first drink and the first incidence of adult-onset DSM-IV alcohol use disorders. *Alcohol Clin Exp Res* 2008;32:2149–60.
- 73 Tanihara S, Momose Y. Reasons for smoking cessation attempts among Japanese male smokers vary by nicotine dependence level: a cross-sectional study after the 2010 tobacco Tax increase. *BMJ Open* 2015;5:e006658
- 74 Freitas SRS, Alvim RO. Smoking and blood pressure phenotypes: new perspective for an old problem. *Am J Hypertens* 2017;30:554–5.
- 75 Saikia B, Marbaniang SP, Kumar P. Changing pattern of tobacco consumption and quitting behavior in northeast India. *J Subst Use* 2021:1–12.
- 76 Wilsnack RW, Wilsnack SC, Kristjanson AF, *et al.* Gender and alcohol consumption: patterns from the multinational GENACIS project. *Addiction* 2009;104:1487–500.
- 77 Nolen-Hoeksema S. Gender differences in risk factors and consequences for alcohol use and problems. *Clin Psychol Rev* 2004;24:981–1010.
- 78 Rasiah R, Yusoff K, Mohammadreza A, *et al.* Cardiovascular disease risk factors and socioeconomic variables in a nation undergoing epidemiologic transition. *BMC Public Health* 2013;13:886.
- 79 Winkleby MA, Jatulis DE, Frank E, *et al.* Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am J Public Health* 1992;82:816–20.
- 80 Menasco M, Blair S. Gender differences in substance use across marital statuses. *Int J Criminol Sociol* 2016;5:1–13.
- 81 Power C, Rodgers B, Hope S. Heavy alcohol consumption and marital status: disentangling the relationship in a national study of young adults. *Addiction* 1999;94:1477–87.