

BMJ Open Patterns of health-related behaviours among adolescents: a cross-sectional study based on the National Survey of School Health Brazil 2012

Catarina Machado Azeredo,¹ Renata Bertazzi Levy,²
 Maria Fernanda Tourinho Peres,² Paulo Rossi Menezes,² Ricardo Araya³

To cite: Azeredo CM, Levy RB, Peres M F T, *et al.* Patterns of health-related behaviours among adolescents: a cross-sectional study based on the National Survey of School Health Brazil 2012. *BMJ Open* 2016;**6**:e011571. doi:10.1136/bmjopen-2016-011571

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2016-011571>).

Received 18 February 2016
 Revised 2 September 2016
 Accepted 7 September 2016



CrossMark

For numbered affiliations see end of article.

Correspondence to
 Professor Catarina Machado Azeredo;
catarina@famed.ufu.br

ABSTRACT

Objectives: The aim of this study was to analyse the clustering of multiple health-related behaviours among adolescents and describe which socio-demographic characteristics are associated with these patterns.

Design: Cross-sectional study.

Setting: Brazilian schools assessed by the National Survey of School Health (PeNSE, 2012).

Participants: 104 109 Brazilian ninth-grade students from public and private schools (response rate=82.7%).

Methods: Exploratory and confirmatory factor analyses were performed to identify behaviour clustering and linear regression models were used to identify socio-demographic characteristics associated with each one of these behaviour patterns.

Results: We identified a good fit model with three behaviour patterns. The first was labelled 'problem-behaviour' and included aggressive behaviour, alcohol consumption, smoking, drug use and unsafe sex; the second was labelled 'health-compromising diet and sedentary behaviours' and included unhealthy food indicators and sedentary behaviour; and the third was labelled 'health-promoting diet and physical activity' and included healthy food indicators and physical activity. No differences in behaviour patterns were found between genders. The problem-behaviour pattern was associated with male gender, older age, more developed region (socially and economically) and public schools (compared with private). The 'health-compromising diet and sedentary behaviours' pattern was associated with female gender, older age, mothers with higher education level and more developed region. The 'health-promoting diet and physical activity' pattern was associated with male gender and mothers with higher education level.

Conclusions: Three health-related behaviour patterns were found among Brazilian adolescents. Interventions to decrease those negative patterns should take into account how these behaviours cluster together and the individuals most at risk.

Strengths and limitations of this study

- This study uses an impressive, nationally representative sample of Brazilian ninth-grade students, and a high response rate was obtained.
- This study investigated the clustering of a broad range of health-related behaviours, enabling a global understanding of how behaviours cluster in this age group.
- We provided evidence on clustering of behaviours in a middle-income country, adding evidence to the literature mostly from high-income countries.
- The use of a sample of students may introduce some selection bias by excluding those youths who drop out of school and who may be most at risk.
- This study was based on the self-reported behaviours of a single group of respondents (the students) which could result in information bias.

INTRODUCTION

Health-compromising behaviours such as substance use (smoking, alcohol and drugs), unsafe sex, poor diet and sedentary behaviour are usually initiated during adolescence and persist throughout life.^{1 2} These behaviours are modifiable causes of morbidity and mortality and thus, must be tackled as a strategy for global health promotion and prevention.³

Interest in clustering of health-related behaviours has increased in the last two decades, however, some issues remain to be addressed by the literature in this field. First, many studies have used inconsistent terminology, such as 'cluster of behaviours' when using cluster analysis, which actually cluster individuals, not behaviours;⁴ and often have used limited statistical analyses, such as analysis of co-occurrence, which do not provide

any indication of whether the concurrence of health-related behaviours is the result of an association between the behaviours.⁵

Second, most of the research in this field has included only the four main risk factors for non-communicable diseases (smoking, alcohol consumption, inadequate diet and physical inactivity).^{4–8} A few studies have assessed a broader range of health-related behaviours among adolescents, including aggressive and sexual behaviour, and these studies have shown a different number of patterns and different variables loading in each pattern.^{9–12}

Finally, despite the importance of studying these behavioural patterns in a wide range of settings, the studies have mostly been conducted in high-income countries. The factors that shape patterns of behaviours, such as social acceptance, culture, legislation and regulatory frameworks, may vary between high-income, low-income and middle-income countries, therefore it is possible that behaviour patterns are different.^{10 13 14} For instance, in the Brazilian setting, the only study among adolescents in this field was limited to a few behaviours and used analysis of co-occurrence, not assessing how these behaviours clustered in specific patterns.⁶

The use of factor analysis to study how a broad range of behaviours cluster in a nationally representative sample of adolescents in a middle-income country may add evidence to the literature. Factor analysis is a suitable method to assess clustering of behaviours and identify underlying factors associated with these behaviours.¹⁴

It is also important to assess whether the behaviour patterns vary according to gender, since, historically, trends of behaviours have shown variation by gender and recent research suggests a decrease in gender differences.¹⁵ Findings from such analyses could have important implications for public health interventions to address multiple risk behaviours and its consequences.⁵

Thus, the aim of the present study is to analyse the clustering of multiple health-related behaviours in a nationally representative sample of ninth-grade students from Brazil, and describe which socio-demographic characteristics are associated with these patterns.

METHODS

Study population, sampling and data collection

We used data from the National Survey of School Health among ninth-grade students (*Pesquisa Nacional de Saúde do Escolar* (PeNSE) 2012), a cross-sectional study, carried out from April to September 2012. The aim of PeNSE is to assess risk and protective factors for health among adolescents enrolled in ninth grade in public and private schools in Brazil.¹⁶

The PeNSE 2012 sample is representative of the country as a whole, the country's five major geographical areas and the 26 state capitals and Federal District. The sampling framework was the 2010 School Census

database, and the sampling strategy included stratification per cluster and multi-stage selection. The sampling strata were each of the 26 state capitals and the Federal District, and the five major geographical areas. In all state capitals and the Federal District, the primary sampling units (PSUs) were schools, and the secondary sampling units (SSUs) were classrooms. In the set comprising the remaining counties in each of five geographical areas, PSUs were county clusters, SSUs were schools, and tertiary sampling units (TSUs) were classrooms. School selection was proportional to the total number of ninth-year classes, while the classes in each school were chosen by simple random selection. Two classrooms were selected from the schools with three or more ninth-year classrooms, and one classroom was selected from the schools with one or two ninth-year classrooms. All of the students enrolled in the selected classrooms were invited to participate in the study.¹⁶

From 3004 selected schools, 162 were not assessed due to absence of ninth-year classrooms, strikes at the time of data collection or the school board's refusal to participate. Considering the total number of students who attended school during data collection (n=110 873, 84% of all students enrolled), participation refusals (n=1651) and lack of report of gender or age (n=118), the final response rate was 82.7%. Data from 109 104 students attending 2842 schools were used. Further details of the sampling design can be found elsewhere.¹⁶

Students filled out smartphone application questionnaires in their school classrooms during regular school hours. The questionnaire was based on the Global School-Based Student Health Survey,¹⁷ the Youth Risk Behaviour Surveillance System¹⁸ adapted to the Brazilian setting.¹⁹ Questions included socio-demographic characteristics, occupation, diet, body image, physical activity, smoking, use of alcohol and other drugs, support network (family and friends), hygiene habits, mental health, oral health, asthma, sexual behaviour, violence and accidents, and use of healthcare services.

Study variables

Health-related behaviours analysed in the study included involvement in physical fights, fights with guns or other weapons (knives, bottles, etc), bullying behaviour (aggressor/bully), alcohol use, drug use, smoking, sexual behaviour, physical activity and sedentary behaviour, and dietary intake of healthy and unhealthy food indicators (see online supplementary appendix 1).

The following socio-demographic covariates were considered in the analysis to describe behaviour clusterings: gender; age range (13 or less; 14–15 years; 16 or more); mother's educational level (incomplete middle school, complete middle school, complete high-school, complete higher education), school administrative status (public or private) and the geographic regions were categorised as more developed (South, Southeast and Centre-West) and less developed (Northeast and North), regarding social and economic indicators.

Statistical analysis

First, a descriptive analysis of the main variables of interest was undertaken according to gender.

Second, exploratory factor analysis (EFA) was performed to generate patterns of health behaviour that aggregated together.⁴ In factor analysis, only shared variance between variables appears in the solution, and the shared variance is partitioned from its unique variance and error variance.²⁰ Sampling adequacy was assessed using the Kaiser-Meyer-Olkin (KMO) criteria. KMO takes values between 0 and 1 with small values, meaning that overall the variables have too little in common to warrant a factor analysis. A value of KMO=0.83 was obtained, meaning good adequacy.²¹ The EFA analysis was performed on the total sample and also on the sample split according to gender.

Factors (patterns) were extracted using weighted least squares with mean and variance adjusted (WLSMV), an estimation method which does not assume multivariate normality and is therefore more appropriate for handling dichotomous or categorically ordered data.²² Using WMLSV in Mplus, the missing data were handled as a function of observed covariates, assuming that the missingness mechanism is 'missing at random'.²³

The choice of the number of factors to be retained was made based on the scree plot assessment, the 'cleanest' factor structure (with item loadings above 0.30), the fewest possible item cross-loadings, no factors with fewer than three items and a reasonable interpretation of the emerging factors.²⁰ The minimum loading of an item to be included in a factor was 0.30.²⁰ The proportion of the variance explained by each factor was obtained from the eigenvalue of each factor divided by the number of variables.²⁴ The oblique rotation (Geomin) method was used because the factors were expected to have some correlation.

After the factor solution was determined, confirmatory factor analysis (CFA) was performed for the factors retained in EFA and the observed variables with loading above 0.3 as indicators of each factor, in order to reduce the complexity of the model. The model fit was assessed through Comparative Fit Index (CFI>0.90), Tucker Lewis Index (TLI>0.90) and root mean square of approximation (RMSEA<0.05). Following CFA, factor scores (continuous variable) were estimated, through regression analysis, for each one of the factors, to identify an individual's placement within the factors.²⁵ Afterwards, we performed linear regression models to describe socio-demographic characteristics associated with each pattern of behaviour (factor scores) in multiple adjusted models. The linear regression assumptions were met.

Multiple imputation was performed using the *chained equation technique* due to the significant proportion of missing values for the mother's education level (17%, n=18 527), and included all other variables with a smaller proportion of missing values, to create a complete data set. The distribution of the observed data was

used to estimate a set of plausible values for the missing data, including random components in the estimated values.²⁶ We performed analysis using the raw data set and the imputed data set (data not shown), for sensitivity analysis. The results from these analyses did not differ. Therefore, the imputed data set was used to run the linear regression models of factor scores.

The cut-off point for statistical significance was established as $p \leq 0.05$. The EFA and CFA analyses were performed using Mplus V.7.0 and both descriptive and regression analyses were performed using Stata SE V.13. The PeNSE data set included weighting factors according to sampling procedures, and this was not incorporated in the factor analysis (EFA, CFA), but in all other analysis.

Ethical aspects

PeNSE 2012 was approved by the National Commission of Research Ethics (Comissão Nacional de Ética em Pesquisa—Conep, record number 16805). It was performed in accordance with the Declaration of Helsinki and all participants gave their informed consent through self-administered questionnaire (smartphones). Informed consent from the next of kin, caretakers or guardians was not obtained on behalf of the participants because the Brazilian Statute of Children and Adolescents (Law number 8.069; 13 July 1990) provides adolescent's autonomy to take initiatives, such as answering a questionnaire that offers no risk to your health and which has the clear purpose of supporting health policies for this age group. All this consent procedure was approved by the National Commission for Research Ethics. Its database was made publicly available on the IBGE website with no information which could identify subjects.

RESULTS

The majority of the students were aged 14–15 years old (63.9%), female (52.2%) and had mothers with low educational level (41.8% incomplete middle school). Most of the students lived in more developed regions (66.8%), and more than 80% of the students attended public schools (table 1).

When analysing the proportion of engagement in health-related behaviours among males and females, it was possible to see some gender differences (see online supplementary appendices 2 and 3). Males presented a higher proportion of all frequencies of involvement in aggressive behaviour when compared to females (see online supplementary table appendix 1). Males were also more involved in unsafe sex (twice as often as females) (see online supplementary table appendix 2).

There were only small gender differences when substance use and food intake indicators were compared. It is worth noting that the highest frequencies of engagement in risk behaviours represented a small proportion of students.

Table 1 Characteristics of the Brazilian adolescents assessed by the National Survey of Health among ninth-grade students—PeNSE, 2012

Variables	Per cent
Gender	
Female	52.2
Male	47.8
Age	
13 years or less	22.9
14–15 years	63.9
16 years or more	13.2
Maternal education	
Incomplete middle school	41.8
Complete middle school	18.0
Complete high school	29.1
Complete higher education	11.1
Geographic region	
Less developed	33.2
More developed	66.8
Administrative status	
Private	17.2
Public	83.8

Three patterns were retained, explaining nearly 50% of the variance. The first pattern was labelled ‘problem-behaviour’, and included involvement in different forms of aggression (physical, with guns or other weapons and verbal bullying), smoking, alcohol consumption, drug use and unsafe sex. The second pattern was labelled ‘health-compromising diet and sedentary behaviours’ and included sedentary behaviour and all unhealthy diet indicators (fizzy drinks, bagged snacks, fried salty snacks). The third pattern was labelled ‘health-promoting diet and physical activity behaviour’ and included physical activity and all healthy diet indicators (fruit, raw vegetables and cooked vegetables) (table 2). The clustering of behaviours was similar for males and females; therefore, the following analysis, including the CFA model and the regression models, were performed for the whole sample, and adjusted by gender.

It is worth noticing small differences in the models based on gender, that in the male model the indicators of involvement in aggression and drug use had strongest correlations in the first pattern; however, they also were correlated with the second pattern at a lower intensity. In the female model this pattern was repeated but for alcohol consumption rather than the aggression and drug use indicators (table 2).

The CFA model showed good fit adjustment (RMSEA=0.053–90% CI 0.053 to 0.054; CFI=0.95; TLI=0.94) and confirmed the patterns extracted from EFA (figure 1) that the problem-behaviour and health-compromising diet and sedentary behaviours patterns were moderately correlated (0.4), whereas problem-behaviour and health-promoting diet and physical activity behaviour, and health-compromising diet and sedentary behaviours and health-promoting diet and physical activity behaviour were weakly correlated.

Linear regression models for the three patterns of behaviour indicated that the problem-behaviour pattern was associated with male gender, older age, more developed region and public schools. The health-compromising diet and sedentary behaviours behaviour pattern was associated with female gender, older age, mother with a higher education level and more developed region. The health-promoting diet and physical activity behaviour was associated with male gender and mother with a higher education level (table 3).

DISCUSSION

The results suggested that the protective behaviours collected together in one cluster, labelled health-promoting diet and physical activity, and the risk behaviours in two main clusters, labelled problem-behaviour and health-compromising diet and sedentary behaviours. The difference between these two risk-behaviour patterns may lie at the level of their social acceptability. According to the literature, violent behaviour, drug use and smoking are less socially acceptable behaviours among adolescents,⁹ and were captured by the problem-behaviour pattern, while the health-compromising diet and sedentary behaviours pattern brought together behaviours that represent harm to health but are not seen as unacceptable either by students or society, such as unhealthy food consumption and sedentary behaviour.⁹ Among Brazilian adults, drinking alcohol and smoking were found as part of a risky lifestyle pattern, together with lack of physical activity and unhealthy diet intake,⁸ in contrast with our findings among adolescents. Among adults, drinking alcohol and smoking are common and socially accepted, while they are contesting or rule-breaking behaviours among adolescents.⁹ This may explain why our study found a problem-behaviour pattern with drinking alcohol and smoking together with aggressive behaviour and not in the compromising behaviour, which included unhealthy diet and sedentary behaviour. In addition, the behaviours comprising the problem-behaviour pattern are more related to establishing independence and gaining peer social acceptance than the behaviours in the health-compromising diet and sedentary behaviours pattern.^{27 28}

In the Netherlands, researchers also found two patterns of risky behaviours among adolescents. These patterns differed from ours regarding the variables that clustered together. Alcohol consumption, unsafe sex and vigorous physical activity clustered together in one pattern, and violent behaviour, smoking and drug use in another.¹⁰ However, these two risky patterns were moderately correlated, as in our study, suggesting that they are not completely independent, and that there are similarities in behaviours clustering among adolescents from these two different settings.

The health-compromising diet and sedentary behaviours pattern was not correlated to the health-promoting diet and physical activity pattern (even negatively),

Table 2 Factor structure of health risk and protective behaviours among Brazilian adolescents

Behaviours	Male			Female			Total		
	Problem-behaviour Factor 1	Health-compromising diet and sedentary behaviours Factor 2	Health-promoting diet and physical activity Factor 3	Problem-behaviour Factor 1	Health-compromising diet and sedentary behaviours Factor 2	Health-promoting diet and physical activity Factor 3	Problem-behaviour Factor 1	Health-compromising diet and sedentary behaviours Factor 2	Health-promoting diet and physical activity Factor 3
Bullying	0.423	0.260	0.030	0.392	0.281	0.052	0.409	0.217	0.009
Physical fight	0.534	0.346	-0.039	0.521	0.262	-0.001	0.532	0.192	-0.100
Fight with gun	0.692	0.406	-0.024	0.590	0.183	-0.034	0.665	0.101	-0.185
Fight with weapon	0.721	0.432	-0.021	0.648	0.226	-0.019	0.704	0.126	-0.179
Smoking—last 30 days	0.747	0.291	0.081	0.743	0.230	0.074	0.739	0.216	0.056
Drinking alcohol—last 30 days	0.914	0.196	-0.027	0.923	0.306	-0.015	0.904	0.374	0.138
Glasses of alcohol	0.906	0.183	-0.015	0.920	0.299	-0.006	0.897	0.366	0.147
Drug use—last 30 days	0.791	0.343	0.062	0.796	0.286	0.061	0.786	0.239	0.014
Number of sexual partners	0.583	0.256	-0.047	0.642	0.095	0.119	0.594	0.127	-0.040
Unsafe sex (without condom)	0.364	0.110	0.131	0.606	0.010	0.213	0.482	0.014	0.088
Bagged salty snacks intake	0.251	0.619	-0.116	0.203	0.659	-0.056	0.115	0.650	-0.049
Fried salty snacks intake	0.258	0.633	-0.184	0.187	0.645	-0.124	0.101	0.641	-0.116
Fizzy drinks intake	0.270	0.542	-0.072	0.223	0.562	-0.005	0.164	0.550	-0.012
sedentary behaviour	0.163	0.266	0.022	0.173	0.323	0.047	0.118	0.302	0.061
Raw vegetables intake	-0.036	-0.039	0.665	0.022	-0.059	0.662	0.068	-0.077	0.632
Cooked vegetables intake	0.001	-0.047	0.609	0.064	-0.025	0.605	0.096	-0.049	0.604
Fruit intake	-0.041	-0.130	0.591	0.041	-0.124	0.589	0.073	-0.136	0.581
Physical activity	-0.095	-0.092	0.305	-0.003	-0.042	0.339	-0.064	-0.035	0.322
Eigenvalue	5280	2067	1508	5222	2050	1667	5273	2027	1640
Explained variance (%)	29.33	11.48	8.38	29.01	11.39	9.26	29.29	11.26	9.11
Cumulative explained variance (%)	29.33	40.81	49.19	29.01	40.4	49.66	29.29	40.55	49.66

National Survey of Health among ninth-grade students—PeNSE, 2012. Bold data reflect that the variable has loaded in the pattern.

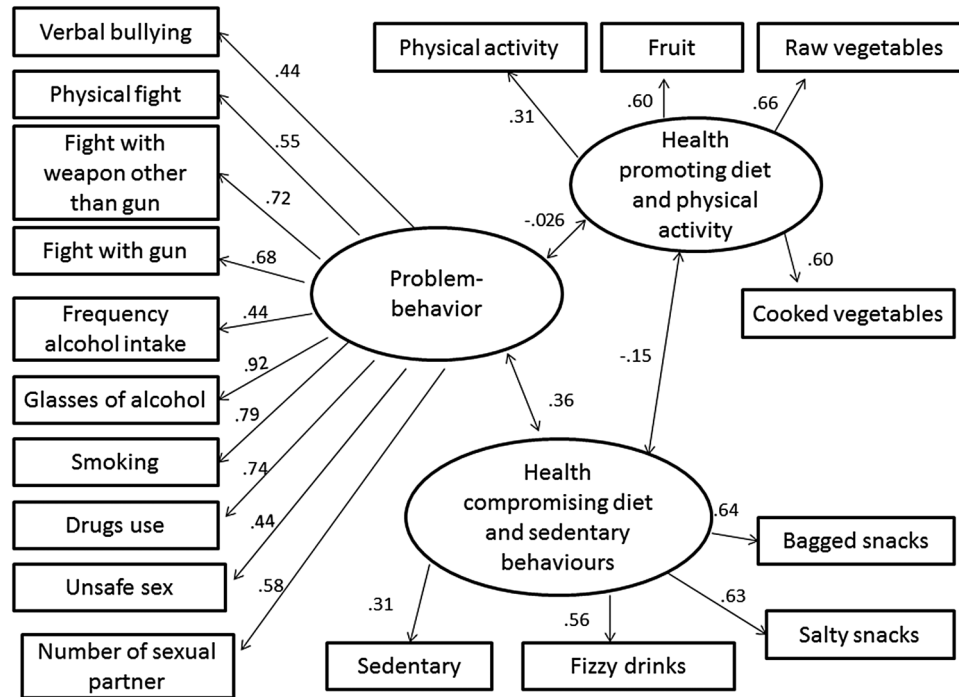


Figure 1 Behaviour patterns and factorial loads assessed by confirmatory factor analysis. National Survey of Health among ninth-grade students—PeNSE, 2012.

Table 3 Association of the three patterns of behaviours score with socio-demographic variables among Brazilian adolescents, PeNSE 2012

Variables	Problem-behaviour Coefficient (95% CI)	Health-compromising diet and sedentary behaviours Coefficient (95% CI)	Health-promoting diet and physical activity Coefficient (95% CI)
Gender			
Female	0	0	0
Male	0.11** (0.09 to 0.14)	-0.07** (-0.11 to -0.03)	0.06* (0.02 to 0.10)
Age			
13 years or less	0	0	0
14–15 years	0.24** (0.22 to 0.27)	0.09** (0.07 to 0.12)	-0.02 (-0.06 to 0.03)
16 years or more	0.63** (0.60 to 0.66)	0.16** (0.08 to 0.24)	-0.04 (-0.09 to 0.00)
Maternal education			
Incomplete middle school	0	0	0
Complete middle school	0.03 (-0.01 to 0.07)	0.13** (0.09 to 0.18)	0.10** (0.07 to 0.12)
Complete high school	0.04 (-0.00 to 0.07)	0.17** (0.14 to 0.21)	0.12** (0.09 to 0.14)
Complete higher education	0.07 (0.01 to 0.14)	0.20** (0.17 to 0.24)	0.25** (0.21 to 0.29)
Geographic region			
Less developed	0	0	0
More developed	0.22** (0.13 to 0.32)	0.23** (0.15 to 0.31)	0.09 (-0.03 to 0.22)
Administrative status			
Private	0	0	0
Public	0.07** (0.03 to 0.11)	-0.06 (-0.14 to 0.02)	0.05 (-0.02 to 0.12)
Constant	-0.052 (-0.56 to -0.48)	-0.20 (-0.28 to -0.012)	-0.12 (-0.23 to -0.01)

* $p < 0.01$; ** $p < 0.001$.

Bold data reflect statistical significance ($p < .05$).

indicating that to engage in a risky lifestyle was determined by a different construct than, for example, not engaging in a healthy lifestyle, possibly due to different predictors. The health-promoting diet and physical

activity pattern may be more influenced by family behaviours, since their components are commonly shared and promoted by family according to the literature.²⁹ On the other hand, the health-compromising diet and

sedentary behaviours pattern is composed of behaviours that may or may not be shared by family and may be more under adolescents' own control, being adopted without their family supervision (in schools, for example).^{30 31} These findings suggest that promoting health-promoting diet and physical activity pattern may not affect the adoption of the health-compromising diet and sedentary behaviours pattern among adolescents. Therefore, these two patterns must be tackled complementarily.

Assessing behaviours individually, we found small differences between genders and the same behaviour patterns for males and females. Other studies have found no gender differences in behaviours, such as alcohol consumption, smoking and drug use, or a decrease in the gap between males and females in Brazil as well as in other parts of the world.^{15 32–34}

In terms of the characteristics of individuals associated with the three patterns, some similarities and differences were found. Older students (>16 years old) had a higher score in the problem-behaviour pattern than younger students. Since 16 years is older than usual among ninth-grade students, it is possible that these students were kept in ninth grade for an extra year due to poor school performance. Older age is also associated with health-compromising diet and sedentary behaviours pattern, but the strength of this association was not as marked. Similar associations with age have been reported before and suggest that prevention programmes may need to focus more closely on those students possibly struggling with school before they engage in multiple risk behaviours.³⁵

Problem-behaviour and health-compromising diet and sedentary behaviours patterns were associated with more developed regions in Brazil, and problem-behaviour was also more frequent in public schools. More developed geographical regions in Brazil are also more urbanised,³⁶ which has been linked to more risky lifestyles in middle-income and low-income countries.³⁷

The health-compromising diet and sedentary behaviours pattern, and health-promoting diet and physical activity pattern were associated with higher maternal education. In this study, maternal education was the only socioeconomic status indicator available. The association found is not surprising, since opportunities to engage in physical activities and access to both healthy and unhealthy food seem to depend on socioeconomic status, something that has been described in relation to food acquisition patterns among Brazilians.³⁸

This study has some limitations. Among these it is the use of a sample of students only. Even though access to basic education is widespread in Brazil (97% and 88% of the population aged 6–14 and 15–19 years old, respectively),³⁹ this may introduce some selection bias by excluding those youths who drop out of school and who may be most at risk.¹³ Accordingly, our results are representative of ninth-grade Brazilian students that regularly attend school. In addition, despite the high

response rate obtained, which reduces the possibility of selection bias, it is still possible that students that had refused to participate or were not at school during the data collection were also at higher risk for the factors assessed. The set of protective behaviours was limited only to diet and physical activity and if more behaviours had been included, such as sleeping habits and hygiene, it is possible more than one underlying factor would have been found. Nevertheless, these two behaviours have been described as the most important to promote health, considering the global burden of disease.⁴⁰

This study was based on the self-reported behaviours of a single group of respondents (the students) which could result in information bias, with overestimation of health-promoting behaviours and underestimation of health-risky behaviours. However, as the students were advised that the study was anonymous and they answered the questions on a smartphone, not through an interview, it is unlikely that this was the case.

We also acknowledge that due to the nature of the data set (students nested within schools), the observations are not independent and, therefore, multilevel analysis could be a valuable alternative when describing the pattern scores. However, multilevel analysis using Stata does not fit sample weights used in the sampling procedure, which could bias the results. On the other hand, the SVY prefix on Stata software, which we have used in multiple linear regression takes into consideration the sample structure (strata, sample units and sample weights) when calculating the estimates. We chose to use multiple linear regression with the SVY prefix instead of multilevel analysis because the clustering effect we found was small and we did not have a contextual hypothesis to test.

The strengths of this study are the large and nationally representative sample of Brazilian adolescents in a middle-income country. In addition, a broad variety of behaviours has been included, which is important for a global understanding of how behaviours cluster in this age group. The use of categorical behaviour variables, instead of dichotomous variables avoided the loss of information about individual differences,⁴¹ and the use of factor analysis that suits these kind of variables produced more reliable results.²²

Our results show that health behaviours tend to cluster among Brazilian adolescents, with these 18 behaviours grouping into three patterns. Interventions regarding health promotion and disease prevention should be designed focusing on behaviour patterns instead of single behaviours, as is often the case. High risk behaviours tend to cluster in the same individuals and it seems rather inefficient to design programmes addressing single unhealthy behaviours. The associations between socio-demographic characteristics and behaviour patterns suggested that older students from more developed regions were the most vulnerable to the health-compromising behaviours patterns. Older students in ninth grade are probably those experiencing

difficulties at school and have been left behind. Therefore, interventions could target this group to tackle multiple health risk behaviours. On the other hand, preventive strategies should be directed to students at an early age.

Future research should also take a step further in this field in trying to understand the mechanisms that give rise to health behaviour clustering, together with their implications for interventions.³

Author affiliations

¹School of Medicine—Federal University of Uberlândia (Faculdade de Medicina—Universidade Federal de Uberlândia), Uberlândia, Brazil

²School of Medicine—University of Sao Paulo (Faculdade de Medicina—Universidade de São Paulo), Sao Paulo, Brazil

³Department of Population Health, London School of Hygiene and Tropical Medicine, Centre for Global Mental Health, London, UK

Twitter Follow Maria Fernanda Peres at @none

Acknowledgements We would like to thank Professor Bianca DeStavola and Camila Aparecida Borges for providing guidance in the exploratory and confirmatory factor analysis.

Contributors CMA, RA, RBL, MFTP and PRM conceptualised the study. CMA, RA and RBL performed statistical analysis. CMA drafted the manuscript. All authors participated in the interpretation of the results and revised critically the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by the Brazilian National Council of Scientific and Technological Development (Centro Nacional de Desenvolvimento Científico e Tecnológico—CNPq) no. 473502/2013-5 awarded to PRM.

Competing interests None declared.

Ethics approval National Commission of Research Ethics—Conep, record no. 16,805.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The original PeNSE data set is publicly available in: <http://www.ibge.gov.br/home/estatistica/populacao/pense/2012/>

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 and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

- Harris KM, Gordon-Larsen P, Chantala K, *et al.* Longitudinal trends in race/ethnic disparities in leading health indicators from adolescence to young adulthood. *Arch Pediatr Adolesc Med* 2006;160:74–81.
- Jackson CA, Henderson M, Frank JW, *et al.* An overview of prevention of multiple risk behaviour in adolescence and young adulthood. *J Public Health (Oxf)* 2012;34(Suppl 1):i31–40.
- Spring B, Moller AC, Coons MJ. Multiple health behaviours: overview and implications. *J Public Health (Oxf)* 2012;34(Suppl 1):i3–i10.
- Hofstetter H, Dusseldorp E, van Empelen P, *et al.* A primer on the use of cluster analysis or factor analysis to assess co-occurrence of risk behaviors. *Prev Med* 2014;67:141–6.
- McAloney K, Graham H, Law C, *et al.* A scoping review of statistical approaches to the analysis of multiple health-related behaviours. *Prev Med* 2013;56:365–71.
- Dumith SC, Muniz LC, Tassitano RM, *et al.* Clustering of risk factors for chronic diseases among adolescents from Southern Brazil. *Prev Med* 2012;54:393–6.
- Li Y, Zhang M, Jiang, *et al.* Co-variations and clustering of chronic disease behavioral risk factors in China: China chronic disease and risk factor surveillance, 2007. *PLoS ONE* 2012;7:e33881.
- Steele EM, Claro RM, Monteiro CA. Behavioural patterns of protective and risk factors for non-communicable diseases in Brazil. *Public Health Nutr* 2014;17:369–75.
- León JC, Carmona J, García P. Health-risk behaviors in adolescents as indicators of unconventional lifestyles. *J Adolesc* 2010;33:663–71.
- van Nieuwenhuijzen M, Junger M, Velderman MK, *et al.* Clustering of health-compromising behavior and delinquency in adolescents and adults in the Dutch population. *Prev Med* 2009;48:572–8.
- Busch V, Van Stel HF, Schrijvers AJ, *et al.* Clustering of health-related behaviors, health outcomes and demographics in Dutch adolescents: a cross-sectional study. *BMC Public Health* 2013;13:1118.
- Neumark-Sztainer D, Story M, Toporoff E, *et al.* Covariations of eating behaviors with other health-related behaviors among adolescents. *J Adolesc Health* 1997;20:450–8.
- Tursz A. Problems in conceptualizing adolescent risk behaviors: International comparisons. *J Adolesc Health* 1997;21:116–27.
- Jessor R. Risk behavior in adolescence: a psychosocial framework for understanding and action. *J Adolesc Health* 1991;12:597–605.
- MacArthur GJ, Smith MC, Melotti R, *et al.* Patterns of alcohol use and multiple risk behaviour by gender during early and late adolescence: the ALSPAC cohort. *J Public Health (Oxf)* 2012;34(Suppl 1):i20–30.
- IBGE. Pesquisa Nacional de Saúde do Escolar 2012. In: Estatística IBGe, ed. Rio de Janeiro, 2013:256.
- World Health Organization (WHO): Global School-Based Student Health Surveillance (GSHS). 2009. <http://www.who.int/chp/gshs/en/>.
- Eaton DK, Kann L, Kinchen S, *et al.* Youth Risk Behavior Surveillance—United States—2009. In: Services DOHAH, ed. USA: Centers for Disease Control and Prevention (CDC), 2010:148. http://www.ibge.gov.br/home/estatistica/populacao/pense/2012/quest_pense_aluno_2012.pdf
- Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research and Evaluation* 2005;10:1–9.
- Hutcheson GD, Sofroniou N. *The multivariate social scientist: Introductory statistics using generalized linear models*. London: Sage Publications, 1999.
- Muthèn BO. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika* 1984;49:115–32.
- Muthèn L, Muthèn B. *Mplus User's Guide*. Los Angeles, CA: Muthèn & Muthèn, 1998–2007.
- Pedhazur E, Pedhazur SL. *Exploratory factor analysis. Measurement, design, and analysis: an integrated approach*. London: Lawrence Erlbaum Associates, 1991:590–630.
- DiStefano C, Zhu M, Mindrila D. Understanding and using factor scores: considerations for the applied researcher. *Practical Assessment Research & Evaluation* 2009;14(20):1–8.
- White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice.
- Schenker M, Minayo M. Risk and protective factors and drug use among adolescence. *Ciência e Saúde Coletiva* 2005;10:707–17.
- Ferguson CJ, Meehan DC. With friends like these...: peer delinquency influences across age cohorts on smoking, alcohol and illegal substance use. *Eur Psychiatry* 2011;26:6–12.
- Videon TM, Manning CK. Influences on adolescent eating patterns: the importance of family meals. *J Adolesc Health* 2003;32:365–73.
- Wouters EJ, Larsen JK, Kremers SP, *et al.* Peer influence on snacking behavior in adolescence. *Appetite* 2010;55:11–17.
- Wills TA, Isasi CR, Mendoza D, *et al.* Self-control constructs related to measures of dietary intake and physical activity in adolescents. *J Adolesc Health* 2007;41:551–8.
- Vieira PC, Aerts DRGdC, Freddo SL, *et al.* Alcohol, tobacco, and other drug use by teenage students in a city in Southern Brazil. *Cad Saude Publica* 2008;24:2487–98.
- Colell E, Sánchez-Niubò A, Domingo-Salvany A. Sex differences in the cumulative incidence of substance use by birth cohort. *Int J Drug Policy* 2013;24:319–25.
- Horta R, Horta B, Pinheiro R, *et al.* Tabaco, álcool e outras drogas entre adolescentes em Pelotas, Rio Grande do Sul, Brasil: uma perspectiva de gênero. *Cad Saude Publica* 2007;23:775–83.
- Neumark-Sztainer D, Story M, French S, *et al.* Patterns of health-compromising behaviors among Minnesota adolescents:

- sociodemographic variations. *Am J Public Health* 1996;86:1599–606.
36. IBGE. Urbanização. In: Atlas do Censo Demográfico 2010., ed. IBGE, Rio de Janeiro, 2013:156.
 37. Popkin BM. Urbanization, lifestyle changes and the nutrition transition. *World Development* 1999;27:1905–16.
 38. Marchioni DM, Claro RM, Levy RB, *et al.* Patterns of food acquisition in Brazilian households and associated factors: a population-based survey. *Public Health Nutr* 2011;14:1586–92.
 39. PNAD. Pesquisa Nacional por Amostra de Domicílio [National Household Sample Survey] 2008/2011, 29/31. Rio de Janeiro: IBGE; 2012.
 40. Schmidt MI, Duncan BB, Silva GA, *et al.* Chronic non-communicable diseases in Brazil: burden and current challenges. *Lancet* 2011;377:1949–61.
 41. MacCallum RC, Zhang S, Preacher KJ, *et al.* On the practice of dichotomization of quantitative variables. *Psychol Methods* 2002;7:19–40.