Ciência & Saúde Coletiva
cienciaesaudecoletiva.com.br
ISSN 1413-8123. v.29, n.1
DOI: 10.1590/1413-81232024291.14712022

Food patterns associated with overweight in 7-11-year old children: machine-learning approach

Padrões alimentares associados ao excesso de peso em crianças de 7 a 11 anos: abordagem de aprendizado de máquina

Emil Kupek (https://orcid.org/0000-0001-6704-1673) [1]
Rafaela Liberali (https://orcid.org/0000-0002-9471-7648) [2]

Abstract *Longitudinal study, whose objective was to present a better strategy and statistical methods, and demonstrate its use with the data across the 2013-2015 period in schoolchildren aged 7 to 11 years, covered with the same food questionnaire (WebCAAFE) survey in Florianopolis, southern Brazil. Six meals/snacks and 32 foods/beverages yielded 192 possible combinations denominated meal/snack-Specific Food/beverage item (MSFIs). LASSO algorithm (LASSO-logistic regression) was used to determine the MSFIs predictive of overweight/obesity, and then binary (logistic) regression was used to further analyze a subset of these variables. Late breakfast, lunch and dinner were all associated with increased overweight/obesity risk, as was an anticipated lunch. Time-of-day or meal-tagged food/beverage intake result in large number of variables whose predictive patterns regarding weight status can be analyzed by machine learning such as LASSO, which in turn may identify the patterns not amenable to other popular statistical methods such as binary logistic regression.*

Key words *Diet surveys, Machine learning, Overweight, Obesity, Childhood*

Resumo *Estudo longitudinal cujo objetivo foi apresentar melhores estratégia e métodos estatísticos e demonstrar sua utilização com os dados do período 2013-2015 em escolares de 7 a 11 anos, contemplados com o mesmo questionário alimentar (WebCAAFE) em Florianópolis, Sul do Brasil. Seis refeições/lanches e 32 alimentos/bebidas resultaram em 192 combinações possíveis denominadas item refeição/lanche-alimentos/bebidas específicos (MSFIs). O algoritmo LASSO (LASSO-regressão logística) foi usado para determinar os MSFIs preditivos de sobrepeso/obesidade e, em seguida, a regressão binária (logística) foi usada para analisar um subconjunto dessas variáveis. Café da manhã, almoço e jantar tardios foram todos associados ao aumento do risco de sobrepeso/obesidade, assim como um almoço antecipado. O consumo de alimentos/bebidas marcados na hora do dia ou na refeição resulta em um grande número de variáveis cujos padrões preditivos em relação ao status do peso podem ser analisados por LASSO. Essa análise pode identificar os padrões não passíveis de outros métodos estatísticos populares, como a regressão logística binária.*

Palavras-chave *Pesquisas de dieta, Aprendizado de máquina, Sobrepeso, Obesidade, Infância*

[1] Departamento de Saúde Pública, Universidade Federal de Santa Catarina. Florianópolis SC Brasil. emil.kupek@ufsc.br
[2] Programa de Póa-Graduação em Ciências Médicas, Universidade Federal de Santa Catarina. Florianópolis SC Brasil.

Introduction

Most food questionnaires applied at early school age are limited to asking about the frequency of consuming foods and beverages of interest, and avoid the questions on their quantity to reduce the cognitive burden of respondents. The studies based on food frequency questionnaires (FFQ) provided only sparse evidence of the association between meal/snack frequency and overweight or obesity[1]. On the other hand, a nationally representative UK study found that boys ate larger portions in the later part of the day[2].

Furthermore, two recent reviews of dietary patterns derived by dimension reduction techniques for school-age children[3,4] found that only a few were predictive of overweight, especially when adjusted for well-established factors that contribute to this outcome, such as child age, sex, family income, and educational level. Although the reviews found high consistency in identifying specific food/beverage items associated with overweight/obesity, no evidence of relevant associations could not be verified through meta-analysis because of the heterogeneity of the studies.

There are few longitudinal studies with children of primary school age. Various studies showed that eating behavior at a later age is influenced by the food intake patterns acquired in the early school years, so that about half of obese children become obese adults[5-7]. Furthermore, dietary patterns of macronutrients are considered a better representation of complex interactions between nutrients[8], thus highlighting the need for longitudinal studies of schoolchildren's dietary patterns.

Virtually all publications on food patterns' association with overweight first derive the patterns by a dimension-reduction method, such as principal components, and factor/cluster/latent class analysis, then relate these patterns to the outcome through logistic or multinomial regression[3,4]. The latter is applied when overweight includes two categories (with and without obesity), so that the weight status has three levels, and normal weight is a natural reference category.

However, this strategy is not optimal if the predictive value of eating patterns is the primary goal of the analysis. The present study aims to present a better strategy and statistical methods to respond to this question and demonstrate its use with the data from a longitudinal study of 7-11-year old children in southern Brazil.

Material and methods

The data were collected in a longitudinal population study of 7-11-year-old children, recruited in 2013, and followed up in 2014 and 2015 with the same WebCAAFE (acronym for "web-based food intake and physical activity evaluation" in Portuguese) questionnaire[9]. Only public schools were targeted and about 95% of these fully participated in the data collection. Three schools without a computer room were excluded. Primary sampling units were 2nd to 5th-grade classrooms, all of whom participated in the surveys. Mentally handicapped and visually impaired children did not take part in this research. Both child and parental consent were obtained for 81.6% of the children.

Among the children with informed consent and complete information on age, sex, weight, height, and food consumption, 9.1% were excluded because of implausible dietary data, such as reporting less than four food items per day or out of the mean ± 3 standard deviation interval.

The studies were conducted according to the guidelines set out in the Code of Ethics of the World Medical Association (Declaration of Helsinki) and all procedures involving human subjects were approved by the Human Studies Committee of the Federal University of Santa Catarina (protocols 037/02, 028/06). Written informed consent was obtained from the parents and oral assent was obtained from the children.

WebCAAFE questionnaire included 32 food and beverage items (bread/biscuits, chocolate milk, coffee with milk, milk, yogurt, cheese, rice, beans, pasta, beef/poultry, fish/seafood, leafy vegetables, cooked vegetables, vegetable soup, fruits, fruit juices, French fries, pizza/hamburger, sweets, salty snacks, and soft drinks) for six meals/snacks: breakfast, morning snack, lunch, afternoon lunch, dinner, night snack. The multiplication of all meals/snacks and foods/beverages combinations yield 192 (6 x 32) items, abbreviated as MSFIs (Meal/snack-Specific Food/beverage Item).

Schoolteachers instructed the children to click on the WebCAAFE icon on the computer screen and start answering the questions; they also made themselves available to clarify possible doubts. An animated robot-like Avatar guided the children while answering the questionnaire. Before closing the block on food consumption, the children were presented with a tray of the foods/beverages they selected for each meal and asked to check and revise their answers if necessary.

Being overweight or obese was the primary outcome of interest. It was defined as a body mass index (BMI) z-score of 2.00 or higher, adjusted for child sex and age[10], following the definition of the World Health Organization. Anthropometric weight and height measurements were obtained by trained physical education teachers following standard procedures[11] from the children who were present at the school on the day of data collection.

Statistical analysis used three steps to evaluate the additional predictive value of food/beverage items and their patterns concerning overweight/obesity as a binary dependent variable, in addition to already established factors predictive of this outcome available from the WebCAAFE: child age, sex, school shift, and the survey year. In the first step, the lasso (least absolute selection and shrinkage operator) based on accuracy improvement in five-fold cross-validation, was used to select the variables associated with the outcome over and above the control variables (child age, sex, school shift, and survey year). In the second step, the lasso-selected and control variables were independent variables for overweight/obesity in multiple logistic regression, and the time-of-day food patterns were extracted from the results, such as drinking milk for breakfast and after dinner, eating sweets only for dinner, and similar. In the third step, post hoc contrasts were calculated to compare the children with a particular food pattern to all other children.

Family income was available only through the mean census sector income depending on the school location and was transformed into quintiles of the income distributions. The choice of the control variables was based on the publications in the area of schoolchild nutrition, in particular with earlier research of the same population[12]. In total, 192 exposure variables (MSFI) and five control variables (child age, sex, school shift, survey year, and family income) were considered for statistical analysis.

Descriptive statistics presented key demographic characteristics and the most frequent MSFIs choices. The commando "xpologit" with the accuracy improvement based on five-fold cross-validation was used within Stata software, version 16.1 (StataCorp. Stata: Release 16. Statistical Software. College Station, TX: StataCorp LLC, 2019). The area under the ROC curve (AUROC) for logistic model accuracy was calculated using a five-fold cross-validation implemented in the cvauroc Stata program.

WebCAAFE has been amply validated and showed good external validity[13,14], reproducibility[15,16], and usability[9]. Its accuracy was shown to be close to that of other similar instruments[13].

The appeal of machine learning methods, such as lasso, to select independent variables predictive of overweight among 192 MSFI variables, lies in their better statistical properties over more traditional methods such as stepwise, especially when there is a large number of independent variables, possibly highly correlated. Stepwise is known as a "greedy" algorithm with a tendency to select too many independent variables, thus overfitting the regression model, whereas lasso is less prone to this difficulty[17]. While the stepwise method produces an inconsistent estimator whose asymptotic distribution is not normal in this situation and does not converge to the true value[18], the lasso selection is consistent[19,20] and has better finite-sample properties, especially when combined with cross-validation[21]. The latter was employed in the present study by partitioning the data into five non-overlapping subsets and averaging over these to calculate the model estimates.

Results

Approximately 3.18% of the WebCAAFE reports over the 2013-2015 period were excluded from analysis due to implausibly low (< 4) total frequency of foods and beverages reported for 24 h, resulting in the analytical sample size of 6585 (Table 1). There were fewer children at the extremes of the age distribution. Also, only a small fraction of the children went to all-day school, and only 1.42% were underweight. Overweight (including obesity) was found in about 35% of the participants.

About 87% of the children consumed breakfast regularly. A typical breakfast (Table 2) consisted of bread/biscuits (45%) and milk with coffee (24%) or chocolate (21%). Less frequently, it also included cream biscuits (11%), milk (10%), and yogurt (8%).

Morning snack was consumed by 63% of the children and typically included bread/biscuits (15%), fruits (12%), yogurt (10%), cream biscuits (9%), and fruit juice (6%) (Table 2). Almost 96% of the children reported having lunch, consisting mainly of rice (57%), cooked beans (44%), chicken/meat (44%), and accompanied by soda drinks (16%) or fruit juice (14%). Afternoon snack was regularly consumed by 81% of the children and predominantly included bread/biscuits (28%), cream biscuits (13%), fruits (12%), yogurt (12%),

Table 1. Socio-demographic characteristics and weight status of 7-11-year-old schoolchildren from public schools in Florianopolis, Brazil.

Characteristic	Description	2013-2015
Sample size obtained		6,801
Analytical sample size		6,585
Excluded from analysis (%)[a]		3.18
Sex	girls	49.25
	boys	50.75
Age (years)	7 (6.50-7.49)	11.29
	8 (7.50-8.49)	22.71
	9 (8.50-9.49)	25.50
	10 (9.50-10.49)	26.12
	11 (10.50-11.49)	14.38
School shift	morning	46.44
	afternoon	51.45
	all day	2.11
Family income quintiles	1st	21.46
	2nd	19.73
	3rd	22.15
	4th	16.73
	5th	19.93
Weight status	Underweight	1.42
	Normal weight	63.58
	Overweight not obese	20.88
	Obese	14.11

[a] due to implausibly low (< 4) total frequency of foods and beverages reported for 24h.

Source: Authors.

chocolate drinks (10%), and coffee with milk (10%) (Table 2).

Dinner was regularly consumed by 90% of the children. Its composition resembled that of the lunch, with rice (34%), beans (25%), chicken/meat (26%), sodas (14%), and fruit juice (13%) being the most frequently consumed items (Table 2). Eating the night snack was reported by 63% of the children, with a menu more varied compared to the other meals, thus making it more difficult to establish a clear food preference. Fruits (9%), sweets (8%), cream biscuits (8%), sodas (7%), fruit juice (7%), and bread/biscuits (7%) were the most frequent choices (Table 2).

Lasso-based odds ratios (OR) for multiple binary logistic regression produced 192 P-values but only those ≤ 0.05 are presented in Table 3.

Eating a traditional Brazilian breakfast including bread/biscuits and milk, possibly with chocolate, was associated with a significant reduction in the risk of being overweight/obese. Reporting consumption of fried food or cream biscuits for the morning snack was also associated with lowering this risk, whereas eating meat or chicken at lunch increased the risk, as did nuggets consumed for the afternoon snack. The children eating sweets or cream biscuits during the latter snack had lower odds of being overweight/obese. Canned foods, eggs, seafood, chips, or similar products consumed for dinner increased the chance of being overweight/obese, whereas eating cooked beans had the opposite effect. The

Table 2. Percentage of top five foods, highlighted in bold letters, consumed for each meal/snack.

Food	Breakfast	Morning snack	Lunch	Afternoon snack	Dinner	Night snack
Rice	2.85	2.29	**57.30**	1.92	**33.94**	2.65
Green leaves	0.81	0.94	12.06	0.61	6.55	0.89
Beans	2.84	2.14	**43.71**	1.42	**25.08**	2.20
Pasta	1.70	2.21	13.27	1.95	11.49	2.32
Chicken/meat	1.87	2.00	**44.26**	2.28	**26.08**	2.56
Fruits	7.74	**12.01**	2.39	12.27	2.87	**8.72**
Bread/biscuit	**45.23**	**15.10**	1.82	27.84	8.13	**6.63**
Cream biscuit	**11.51**	**9.09**	0.89	13.19	1.93	**7.91**
Coffee & milk	**23.63**	4.60	0.55	9.76	2.84	2.59
Milk	**9.67**	2.79	0.47	3.88	1.56	4.35
Yogurt	**8.48**	**10.01**	0.48	11.68	1.34	5.40
Chocolate	**21.15**	5.36	0.70	9.87	2.37	**6.71**
Fruit juice	6.41	**6.24**	**13.96**	8.69	12.66	6.85
Sodas	2.92	3.20	**15.70**	6.22	**14.25**	**7.36**
Sweets	1.54	2.32	1.61	6.07	2.57	**8.41**
Cake without cream	6.88	4.43	0.89	9.15	2.57	4.18

Source: Authors.

Table 3. Lasso-based odds ratios for overweight/obesity with P-value < 0.05 among 192 binary independent food/beverage variables, controlled for age, sex, school shift, family income, and survey year.

Meal/ snack	Food/ beverage	OR, 95%CI limits			P
		OR	lower	upper	
Breakfast	Chocolate milk	0.81	0.71	0.94	0.005
	Milk	0.82	0.68	0.99	0.043
	Bread/ biscuit	0.88	0.78	0.99	0.028
Morning snack	Chips & similar	0.51	0.30	0.89	0.017
	French fries	0.61	0.40	0.94	0.027
	Cream biscuit	0.81	0.67	0.99	0.042
Lunch	Meat/ chicken	1.13	1.00	1.27	0.049
Afternoon snack	Sweets	0.76	0.60	0.97	0.026
	Cream biscuit	0.81	0.69	0.96	0.016
	Nuggets	2.35	1.08	5.14	0.032
Dinner	Cooked beans	0.77	0.65	0.92	0.004
	Canned	1.25	1.02	1.52	0.031
	Eggs	1.32	1.02	1.71	0.035
	Seafood	1.38	1.02	1.88	0.035
	Chips & similar	1.61	1.01	2.57	0.047
Night snack	Cream biscuit	0.79	0.64	0.98	0.028
	Vegetables	1.81	1.00	3.25	0.049

OR = odds ratio, CI = confidence interval.

Source: Authors.

vegetables consumed at the night snack were associated with the increase in the odds of being overweight/obese while eating cream biscuits for this snack pointed to a decrease in these odds.

The lasso-selected variables predictive of overweight/obesity were further analyzed in multiple binary logistic regression to look for time-of-day patterns in consuming specific foods or beverages associated with the outcome (Table 4).

Eating cereals and drinking milk at lunchtime may be indicative of late breakfast as these items are typically consumed in the first-morning meal. A small group of 38 (17 + 21) children with these MSFIs had their risk of being overweight increased by 24.7% and 17%, respectively, compared to those with different MSFIs (Table 4).

Eating green leaves (salad) only after dinner may be due to dieting, and was associated with an additional overweight risk of 15.2% compared to those who did not share this eating pattern. The same magnitude of effect was found for eating nuggets for lunch.

Drinking milk exclusively for dinner may be indicative of dieting, and increased the risk of being overweight by 13.8% in comparison to all other children (Table 4). A slightly lesser magnitude of the risk effect was found for eating cooked beans at morning snacks, suggesting an anticipated lunch, and for consuming vegetable soup at major meals, possibly indicative of dieting.

Table 4. Statistically significant (P < 0.1) average marginal increase in the probability of overweight regarding selected eating patterns, based on post-hoc multivariate logistic regression contrasts (N = 6,420).

Food/ beverage	Daily meals						Interpretation	n	%	DPO (95% CI)	P-value
	B	M	L	A	D	N					
Cereals			X				Late B	17	0.26	24.7 (4.3 - 45.2)	0.018
Milk			X				Late B	21	0.33	17.0 (-1.8 - 35.8)	0.077
Green leaves						X	Late D, dieting?	27	0.42	15.2 (0.3 - 30.2)	0.045
Nuggets			X				Fast food	50	0.78	15.2 (1.0 - 29.4)	0.036
Milk					X		Dieting?	64	1.00	13.8 (1.3 - 26.3)	0.031
Beans		X					Anticipated L	64	1.00	13.6 (1.0 - 26.1)	0.034
Vegetable soup			X		X		Dieting?	63	0.98	13.4 (0.7 - 26.1)	0.038
Instant pasta						X	Late D	102	1.59	12.3 (2.6 - 22.1)	0.013
Sweets			X				L + sweets	72	1.12	11.7 (0.2 - 23.2)	0.047
Canned				X			Fast food, late L	95	1.48	11.3 (1.2 - 21.4)	0.029
Vegetables			X		X		Dieting?	168	2.62	11.2 (3.7 - 8.8)	0.004
Fish/seafood					X		Mainly fried?	194	3.02	8.5 (1.4 - 15.5)	0.019

B = breakfast, M = morning snack, L = lunch, A = afternoon snack, D = dinner, N = night snack (after dinner), n = group size, % = percentage of the total (n/N · 100), DPO = difference in the probability of overweight adjusted for age, sex, survey year, school shift and family income, CI = confidence interval.

Source: Authors.

Eating instant pasta exclusively after dinner may be suggestive of fast food being prepared for late dinner, and added 12.3% of overweight risk compared to all other eating patterns (Table 4). Having sweets on the lunch menu only was associated with a risk increase of 11.7%. Eating canned food as an afternoon snack may be indicative of a fast-food solution for late lunch, with an 11.3% risk increase. A similar risk increase was found for eating vegetables for major meals, possibly associated with dieting. The 8.5% risk increase for eating fish/seafood exclusively at dinner may be due to the local cultural preference for frying these foods, thus adding a significant amount of fat.

In total, 940 (14.6%) of the children belonged to the patterns with increased overweight/obesity risk, with the median risk increase in the probability of 13.5% and interquartile range between 11.5 and 15.2% (Table 4).

Finally, two multiple logistic regression models with overweight as the dependent variable were compared in terms of the AUROC: using five control variables as the only independent predictors versus adding lasso-selected variables. The former model showed an AUROC of 0.55 (95%CI 0.52-0.56), compared to the latter with a significantly improved AUROC of 0.62 (95%CI 0.60-0.63).

Discussion

To the best of the author's knowledge, this is the first study to derive MSFIs predictive of overweight directly by penalized binary (logistic) regression using a lasso. It allowed simultaneous analysis of 192 binary exposure variables and identified 12 eating patterns with an increased risk of being overweight.

A major methodological shift in the present study compared to the other eating pattern studies was to search for the patterns based on their predictive value instead of based on the homogeneity regarding food/beverage intake. The latter needs an additional step to relate the patterns to the outcome of interest, and their extraction is not guided by the power to separate those with versus those without the outcome. Principal component, common factors, cluster, and latent class analyses are the most popular dimension-reduction techniques to identify food intake, physical activity, and other lifestyle patterns[3,4].

However, it is a discriminant analysis that searches for distinct combinations of independent variables regarding the outcome. Binary logistic regression breaks down with hundreds of exposure variables, especially when many of these are low-frequency binary variables. In other words, there is a need for sparse discriminant analysis to identify dietary patterns in this situation[22]. A modest but significant increase in the predictive accuracy of overweight/obesity with the addition of lasso-selected variables points to the potential of this analysis to improve the classification of individuals at risk of excess weight during childhood and early adolescence.

Machine learning algorithms sometimes used to be criticized as a "black box" approach but their utility has proved indispensable in genomics, chemometrics, and other "big data" research, especially when the number of subjects/units relative to the number of variables is small. With time-of-the-day or meal tag, the number of variables easily reaches hundreds, not to mention other lifestyle indicators such as physical activity and screen behavior.

Among many machine learning algorithms such as optimal pruning, bagging, boosting, and neural networks, lasso was chosen for its solid statistical theory as opposed to predominantly computational definitions of some of these algorithms. Moreover, the lasso has been extensively tested and widely applied in many areas of science[17].

Most of the results from the present study are in line with those of other dietary pattern studies[3,4] regarding obesogenic foods, such as high-fat and/or fast-food items (nuggets, instant pasta, fried fish/seafood, and canned food) and dieting (consuming green leaves/salad after dinner, vegetables and vegetable soup for major meals). Postponing meals may result in consuming MSFIs later in the day or skipping them altogether. The present study showed that late breakfast, lunch, and dinner were all associated with increased overweight/obesity risk, as was anticipating lunch. This is in line with other studies that found evidence for the importance of the meal timing, such as chronobiological studies related to diet.

Negative feedback regulates the intake of macronutrients both in the short and long term, with the latter peaking at two days[23]. Eating regular meals and snacks with higher frequency and stable schedule helps regulate weight[24]. In adults, higher energy intake at lunchtime and lower intake in the evening are both associated with a lower risk of overweight/obesity[25].

A similar conclusion can be drawn for schoolchildren whose dietary patterns include energy-dense traditional Brazilian foods, such as rice and cooked beans, consumed at midday[26]. In mice, eating during nighttime disturbed circadian rhythms and led to leptin resistance, physical inactivity, excessive eating, adiposity increase, metabolic disorders, and obesity[27]. The latter points to likely physiological mechanisms in humans too.

The results of a study based on daily food/beverage intake of the same data used in the present study showed no statistically significant association between the dietary patterns extracted by latent class analysis and child weight status[28]. However, when the data were meal-tagged and analyzed with lasso, the association was found for several patterns, thus reinforcing the importance of the time-of-day analysis. The finding of such analysis is important to develop effective interventions toward a healthy diet, that take into account both group-specific (age, sex, income, culture) and time-of-day-specific dietary patterns[28,29].

A notable sparsity of time trend studies on food consumption in early school age makes it difficult to compare the present study findings with those of the other trend analysis, such as similar studies with adolescents. In Brazil, unhealthy eating habits showed a significant increase, especially among adolescents from low-income families[30]. The consumption of ultra-processed foods has increased at the expense of unprocessed foods such as rice, beans, and fruits, although sugary drink intake has decreased in the last decade[31]. A similar decreasing trend in the consumption of sugary drinks was found in the USA among children and adolescents[32,33]. European adolescents have increased their intake of fruits and vegetables over the decade of 2010[34]. In most developed countries, the consumption of dairy products decreases significantly in adolescence compared to childhood[35].

Among the limitations of the present study, the lack of the food/beverage quantity and possible information bias, inherent to all FFQs, should be kept in mind. Also, each child responded about his/her diet for only one 24-hour period, but the representativeness of the response on the population level should have been maintained as the questionnaires applied referred to every day of the week and one weekend day (Sunday). Furthermore, the meaning assigned to the MSFIs was rather qualitative than quantitative, thus prone to misinterpretation, in particular concerning dieting. Also, the values of AUROC obtained in the present study were still rather low for individual classification, and require further investigation and improvement before a routine use in school settings.

The strengths of this study include large coverage (95%) of the target population and large sample size, thus resulting in large statistical power to detect relatively small associations between dietary patterns and overweight/obesity, as well as the use of an amply validated instrument.

Conclusions

Late breakfast, lunch, and dinner were all associated with increased overweight/obesity risk, as was an anticipated lunch. Time-of-day or meal-tagged food/beverage intake results in a large number of variables whose predictive patterns regarding weight status can be analyzed by lasso. Such analysis may identify the patterns not amenable to be found by other popular statistical methods such as binary logistic regression.

Collaborations

All authors were involved in analyzing the studies, reviewing and interpreting the results and writing the manuscript. All authors takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation.

Acknowledgements

We are grateful to the Post-Doctoral National Program-CAPES (PNPD-CAPES) (Programa Nacional de Pós-Doutorado, PNPD-CAPES) that granted a postdoctoral scholarship.

Funding

Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina – FAPESC (grant n. 062/2002), Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (grant n. 402322/2005-3), and the Brazilian Ministério da Saúde (Departamento de Ciência, Tecnologia e Insumos Estratégicos – DECIT) (grant n. 305148/2011-7).

References

1. Murakami K, Livingstone MB. Decreasing the number of small eating occasions (<15 % of total energy intake) regardless of the time of day may be important to improve diet quality but not adiposity: a cross-sectional study in British children and adolescents. *Br J Nutr* 2016; 115(2):332-341.
2. Blundell-Birtill P, Hetherington MM. Determinants of portion size in children and adolescents: insights from the UK National Diet and Nutrition Survey Rolling Programme (2008-2016). *Nutrients* 2019; 11(12):2957.
3. Liberali R, Kupek E, Assis MAA. Dietary patterns and childhood obesity risk: a systematic review. *Child Obes* 2020; 16(2):70-85.
4. Liberali R, Del Castanhel F, Kupek E, Assis MAAD. Latent class analysis of lifestyle risk factors and association with overweight and/or obesity in children and adolescents: systematic review. *Child Obes* 2021; 17(1):2-15.
5. Katzmarzyk PT, Barlow S, Bouchard C, Catalano PM, Hsia DS, Inge TH, Lovelady C, Raynor H, Redman LM, Staiano AE, Spruijt-Metz D, Symonds ME, Vickers M, Wilfley D, Yanovski JA. An evolving scientific basis for the prevention and treatment of pediatric obesity. *Int J Obes* 2014; 38(7):887-905.
6. Yanovski JA. Pediatric obesity. An introduction. *Appetite* 2015; 93:3-12.
7. Rito A, Sousa RCD, Mendes S, Graça P. Childhood Obesity Surveillance Initiative: COSI Portugal. 2016.
8. Johnson L, Toumpakari Z, Papadaki A. Social gradients and physical activity trends in an obesogenic dietary pattern: cross-sectional analysis of the UK National Diet and Nutrition Survey 2008-2014. *Nutrients* 2018; 10(4):388.
9. Costa FF, Schmoelz CP, Davies VF, Di Pietro PF, Kupek E, Assis MA. Assessment of diet and physical activity of Brazilian schoolchildren: usability testing of a web-based questionnaire. *JMIR Res Protoc* 2013; 2(2):e31.
10. Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007; 85(9):660-667.
11. Lohman TG, Roche AF, Martorell R. *Anthropometric standardization reference manual*. Champaign: Human Kinetics; 1988.
12. Leal DB, Assis MAA, Conde WL, Lobo AS, Bellisle F. Individual characteristics and public or private schools predict the body mass index of Brazilian children: a multilevel analysis. *Cad Saude Publica* 2018; 34(5):e00053117.
13. Davies VF, Kupek E, Assis MAA, Natal S, Di Pietro PF, Baranowski T. Validation of a web-based questionnaire to assess the dietary intake of Brazilian children aged 7-10 years. *J Hum Nutr Diet* 2015; 28(Suppl. 1):93S-102S.
14. Kupek E, Assis MA, Bellisle F, Lobo AS. Validity of WebCAAFE questionnaire for assessment of schoolchildren's dietary compliance with Brazilian Food Guidelines. *Public Health Nutr* 2016; 19(13):2347-2356.

15. Assis MA, Kupek E, Guimarães D, Calvo MC, Andrade DF, Bellisle F. Test-retest reliability and external validity of the previous day food questionnaire for 7-10-year-old school children. *Appetite* 2008; 51(1):187-193.
16. Jesus GM, Assis MAA, Kupek E. Validity and reproducibility of an Internet-based questionnaire (Web-CAAFE) to evaluate the food consumption of students aged 7 to 15 years. *Cad Saude Publica* 2017; 33(5):e00163016.
17. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning*. New York: Springer; 2009.
18. Plotscher BM, Leeb H. On the distribution of penalized maximum likelihood estimators: The LASSO, SCAD, and thresholding. *J Multivariate Analysis* 2009; 100(9):2065-2082.
19. Belloni A, Chernozhukov V, Hansen C. High-dimensional methods and inference on structural and treatment effects. *J Econ Perspect* 2014; 28(2):29-50.
20. Belloni A, Chernozhukov V, Hansen C. Inference on treatment effects after selection among high-dimensional controls. *Rev Econ Studies* 2014; 81(2):608-650.
21. Chernozhukov V, Chetverikov D, Demirer M, Duo E, Hansen C, Newey W, Robins J. Double/debiased machine learning for treatment and structural parameters. *Econometrics J* 2018; 21(1):C1-C68.
22. Clemmensen L, Hastie T, Witten D, Ersboll B. Sparse discriminate analysis. *Technometrics* 2011; 53(4):406-413.
23. Castro JM. What are the major correlates of macronutrient selection in Western populations? *Proc Nutr Soc* 1999; 58(4):755-763.
24. Ekmekcioglu C, Touitou Y. Chronobiological aspects of food intake and metabolism and their relevance on energy balance and weight regulation. *Obes Rev* 2011; 12(1):14-25.
25. Wang JB, Patterson RE, Ang A, Emond JA, Shetty N, Arab L. Timing of energy intake during the day is associated with the risk of obesity in adults. *J Hum Nutr Diet* 2014; 27(Suppl. 2):255-262.
26. Kupek E, Lobo AS, Leal DB, Bellisle F, Assis MA. Dietary patterns associated with overweight and obesity among Brazilian schoolchildren: an approach based on the time-of-day of eating events. *Br J Nutr* 2016; 116(11):1954-1965.
27. Yasumoto Y, Hashimoto C, Nakao R, Yamazaki H, Hiroyama H, Nemoto T, Yamamoto S, Sakurai M, Oike H, Wada N, Yoshida-Noro C, Oishi K. Short-term feeding at the wrong time is sufficient to desynchronize peripheral clocks and induce obesity with hyperphagia, physical inactivity and metabolic disorders in mice. *Metabolism* 2016; 65(5):714-727.
28. Lobo AS, Assis MAA, Leal DB, Borgatto AF, Vieira FK, Di Pietro PF, Kupek E. Empirically derived dietary patterns through latent profile analysis among Brazilian children and adolescents from Southern Brazil, 2013-2015. *PLoS One* 2019; 14(1):e0210425.
29. Smith KL, Straker LM, Kerr DA, Smith AJ. Overweight adolescents eat what? And when? Analysis of consumption patterns to guide dietary message development for intervention. *J Hum Nutr Diet* 2015; 8(Suppl. 2):80-93.
30. Haddad MR, Sarti FM. Sociodemographic determinants of health behaviors among Brazilian adolescents: trends in physical activity and food consumption, 2009-2015. *Appetite* 2020; 144:104454.
31. Martins APB, Levy RB, Claro RM, Moubarac JC, Monteriro CA. Increased contribution of ultra-processed food products in the Brazilian diet (1987-2009). *Rev Saude Publica* 2013; 47(4):656-665.
32. Beck AL, Martinez S, Patel AI, Fernandez A. Trends in sugar-sweetened beverage consumption among California children. *Public Health Nutr* 2020; 23(16):2864-2869.
33. Bleich SN, Vercammen KA, Koma JW, Li Z. Trends in beverage consumption among children and adults, 2003-2014. *Obesity* 2018; 26(2):432-441.
34. Vereecken C, Pedersen TP, Ojala K, Krølner R, Dzielska A, Ahluwalia N, Giacchi M, Kelly C. Fruit and vegetable consumption trends among adolescents from 2002 to 2010 in 33 countries. *Eur J Public Health* 2015; 25 (Suppl. 2):16-19.
35. Dror DK, Allen LH. Dairy product intake in children and adolescents in developed countries: trends, nutritional contribution, and a review of association with health outcomes. *Nutr Rev* 2014; 72(2):68-81.

Article submitted 10/10/2022
Approved 21/03/2023
Final version submitted 23/03/2023

Chief editors: Romeu Gomes, Antônio Augusto Moura da Silva