Clustering of diet, physical activity, and health behaviors in adolescents and association with obesity: A latent class analysis

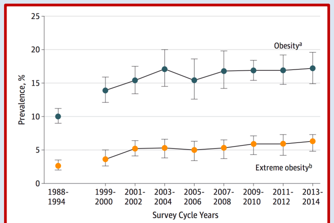
Rebecca Boehm, Ph.D.; Kristen Cooksey Stowers, Ph.D.; Marlene Schwartz, PhD
Rudd Center for Food Policy & Obesity
University of Connecticut
Hartford, CT

UCONN RUDD CENTER FOR FOOD POLICY & OBESITY

UCONN RUDD CENTER FOR FOOD POLICY & OBESITY

Background & Significance

Figure 1. Prevalence of Obesity and Extreme Obesity in U.S. Children and Adolescents Aged 2-19 from 1988-2014.



Source: Ogden et al. 2014, *JAMA*
[a] BMI >95% percentile by age and gender.
[B] BMI ≥120% of the 95% percentile by age and gender.

Childhood obesity is a major public health concern in the U.S. The Centers for Disease Control and Prevention (CDC) report that nearly 17% of children in the U.S. ages 2-19 were obese between 2011 and 2014 (Ogden et al. 2014). Furthermore there has been no major decline in rates of obesity among children in the U.S. over the last 15 years (**Figure 1**), even though interventions and policy changes have been enacted to help children maintain a healthy weight.

The multifaceted nature of childhood obesity is the main reason why it is such an intractable problem. A child's likelihood of being overweight or obese is a function of a whole host of factors such as diet, physical activity, genetics, socioeconomic status, culture, and the safety and quality of their neighborhood and school environments (Nestle et al. 1998; Deckelbaum and Williams, 2001; Dietz and Gortmaker, 2001; Lobstein et al. 2004). Consequently, the specific reason why a particular child is obese is in many ways unique to that child.

This suggests that future interventions and policy changes must take a tailored approach to obesity prevention by accounting for all the ***specific risk factors*** faced by individual children. In addition, health behaviors may interact and compound, dramatically increasing the risk of obesity for a particular child. Therefore, this study examines how, simultaneously, dietary habits, physical activity, socioeconomic status, and other factors interact and are associated with childhood obesity among six grade children in a county in the U.S. We use a Latent Class Analysis (LCA) to identify homogenous groups of adolescents based on self-reported sugary drink and fast food consumption, screen time, social media use, sleep hygiene, and measures of family cohesion.

LCA is increasingly used in the public health field to identify latent or unobservable groups that have similar patterns of risk for a particular disease or adverse health behavior (Collins and Lanza, 2013). In the obesity literature this method and a similar method called cluster analysis have been used to identify the clustering or patterns of behaviors that increase the likelihood of obesity (Dumuid et al. 2016; Leech et al. 2014; Perez Rodrigo, et al. 2015; Fleary et al. 2017; Iannotti and Wang, 2013; Berlin et al. 2017). For example, Laxer et al. (2011) used LCA to identify clustering of health behaviors among Canadian youth and identified four latent groups: "traditional school athletes", "inactive screenagers", "health conscious", and "moderately active substance users". Based on these results the authors suggested for each group a tailored interventional approach for reducing obesity risk and engagement in other unhealthy behaviors. Huh et al. (2011) called ***LCA a "person-centered" empirical approach*** which could inform the development of more effective interventions to combat obesity, especially among high risk populations.

Hypotheses

1. Distinct latent groups of adolescents will exist within our sample based on patterns of behavior related to self-reported consumption of sugary drinks, healthy drinks, and fast food, screen time, social media use, sleep hygiene, and family cohesion.
2. The probability that a child belongs to a particular latent group will be statistically associated with sociodemographic characteristics such as race, gender, and eligibility for free or reduced school meals (a proxy for social class/family income).
3. The probability that a child belongs to a particular latent group will also be statistically associated with their weight status, defined as underweight, normal weight, overweight, or obese using CDC percentile cut points.

Sample and Latent Class Indicators

Data for this study was collected from sixth grade students (n=2,791) attending school in a U.S. county (due to disclosure risk the county cannot be named) during the 2015-2016 school year. Two phases of data collection occurred. First, students completed an electronic survey consisting of questions about consumption of sugary drinks, healthy drinks, and fast food, number of hours of television and computer/electronic device usage per day, average bed time, frequency of breakfast consumption, and frequency of meals consumed with family members. Second, Teachers of physical education (PE) classes measured and recorded each student's height (inches) and weight (pounds) once each school semester for a total of two times for the 2015-2016 school year.

Data collected in the survey was used to construct 13 latent class indicators. Each indicator was chosen because we considered it to be associated with a child's risk of obesity. The indicators (each binary, taking on the value 0 or 1) are reported in **Table 1**. Height and weight measurements during PE classes were used to calculate each student's Body Mass Index (BMI, kg/m^2) and their weight status as determined by percentile cut points by age and sex prescribed by the Centers for Disease Control and Prevention (<5th percentile = underweight, 5th – 95th percentile = healthy weight, 85th to 95th percentile = overweight, >95th percentile = obese).

Statistical Analysis

We utilized the Pennsylvania State University LCA plugin in Stata (version SE, 14.2, StataCorp, Inc.) to run the LCA models with the indicator variables listed in **Table 1**. Models were run sequentially with 1 to 10 latent classes. The best-fitting model was selected based on analysis of various criteria including model efficiency, conceptual accuracy, prediction power, and model efficiency. LCA relies on a series of statistics generated by each model called Information criteria to determine model efficiency. These information criteria statistics are generated based on model fit (i.e. log-likelihood value) and the number of indicators included in the model. Four information criteria were used to assess model fit for this analysis, and they include the: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Consistent Akaike Information Criterion (CAIC), and the adjusted-BIC Criterion (Adj-BIC). After running models with 1-10 classes the AIC values suggested a 10 class model was optimal. However, conceptual accuracy was reduced in the 10 class model compared to models with fewer classes. Further, the BIC, CAIC, and adj-BIC criteria values suggested that a five or six class model was optimal. Class distinctness was readily apparent in the five and six class models, whereas it was not for models with more classes. After comparison of the five and six class models we selected a six class model as having the best fit overall based on conceptual accuracy, prediction power, and model efficiency.

Item response probabilities generated for each dichotomous indicator variable in **Table 1** are used to describe the latent classes identified by the optimal model. Item response probabilities ≥.70 mean that a respondent has a high probability of responding "yes" to the survey question used to generate that indicator variable. Conversely, an item response probability ≤.30 means that a respondent has a low probability of responding "yes" to the survey question used to generate the indicator variable. Indicator variables with item response probabilities ≤.30 and ≥.70 were used to describe each of the classes in the six class model.

The proportion of individuals in each class is not assumed to be equal. For the optimal model selected the actual proportion of individuals assigned to each class is generated by the LCA plugin tool. These proportions or "class shares" are calculated as the mean probability of membership to each class. This tells you how prevalent the class is within the sample of individuals.

Individual class membership probabilities are used to show the association between the probability of class membership for each individual and their sociodemographic characteristics. Individual class membership probability is regressed on sociodemographic characteristics as follows:

$$Y_{i,c} = \beta_0 + \beta_1(gender_i) + \beta_2(race_i) + \beta_3(free/reduced_i) + \varepsilon$$

where Y is the probability of membership to class c for individual i. β parameters determine the association between membership to class c and each characteristic. Characteristics included in the above model are: gender (male, female), race (White, Black, Hispanic, Asian, other race), and whether or not the student is eligible for free or reduced priced school meals. These characteristics are examined because they have been shown in previous studies to be associated with obesity risk in children.

Table 1. Latent Variable Indicators

Category	Indicators
Sugary drinks	Daily consumption of regular soda
	Daily consumption of fruit or sports drinks
	Daily consumption of energy drinks, sweetened flavored water/ tea
Healthy drinks	Daily consumption of 100% juice
	Daily consumption of plain milk (full fat or other types)
Unhealthy foods	Eats fast food ≥2 times per week
	Has a snack at an unhealthy retail outlet (fast food restaurant or corner store ≥2 times per week.
Family cohesion	Has ≥5 meals with family members per week
	Has breakfast ≥5 times per week
Screen time	≥2 hours of TV or non-school work related computer time per day
Social media use	Owns their own smartphone or cellphone
	Uses >1 type of social media (e.g. Twitter, Facebook, Instagram, Snapchat) regularly
Sleep hygiene	Usually goes to sleep after 10pm.

References

Collins, Linda M., and Stephanie T. Lanza. 2013. *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Vol. 718. John Wiley & Sons.

Deckelbaum, Richard J., and Christine L. Williams. 2001. "Childhood Obesity: The Health Issue." *Obesity Research* 9 (S11): 239S–243S. doi:10.1038/oby.2001.125.

Dietz, William H., and Steven L. Gortmaker. 2001. "Preventing Obesity in Children and Adolescents." *Annual Review of Public Health* 22 (1): 337–53. doi:10.1146/annurev.publhealth.22.1.337.

Dumuid, Dorothea, T. Olds, L. K. Lewis, J. A. Martin-Fernández, T. Barreira, S. Broyles, J.-P. Chaput, et al. 2016. "The Adiposity of Children Is Associated with Their Lifestyle Behaviours: A Cluster Analysis of School-Aged Children from 12 Nations." *Pediatric Obesity*, January, n/a-n/a. doi:10.1111/ijpo.12196.

Fleary, Sasha A. 2017. "Combined Patterns of Risk for Problem and Obesogenic Behaviors in Adolescents: A Latent Class Analysis Approach." *Journal of School Health* 87 (3): 182–93. doi:10.1111/josh.12481.

Huh, Jimi, Nathaniel R. Riggs, Donna Spruijt-Metz, Chih-Ping Chou, Zhaoqing Huang, and MaryAnn Pentz. 2011. "Identifying Patterns of Eating and Physical Activity in Children: A Latent Class Analysis of Obesity Risk." *Obesity* 19 (3): 652–58. doi:10.1038/oby.2010.228.

Iannotti, Ronald J., and Jing Wang. 2013. "Patterns of Physical Activity, Sedentary Behavior, and Diet in U.S. Adolescents." *Journal of Adolescent Health* 53 (2): 280–86. doi:10.1016/j.jadohealth.2013.03.007.

Laxer, Rachel E., Ross C. Brownson, Joel A. Dubin, Martin Cooke, Ashok Chaurasia, and Scott T. Leatherdale. 2017. "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* 17: 102. doi:10.1186/s12889-017-4034-0.

Leech, Rebecca M., Sarah A. McNaughton, and Anna Timperio. 2014. "The Clustering of Diet, Physical Activity and Sedentary Behavior in Children and Adolescents: A Review." *International Journal of Behavioral Nutrition and Physical Activity* 11: 4. doi:10.1186/1479-5868-11-4.

Lobstein, T., L. Baur, and R. Uauy. 2004. "Obesity in Children and Young People: A Crisis in Public Health." *Obesity Reviews* 5 (May): 4–85. doi:10.1111/j.1467-789X.2004.00133.x.

Nestle, Marion, Rena Wing, Leann Birch, Lorelei DiSogra, Adam Drewnowski, Suzette Middleton, Madeleine Sigman-Grant, Jeffery Sobal, Mary Winston, and Christina Economos. 1998. "Behavioral and Social Influences on Food Choice." *Nutrition Reviews* 56 (5): 50–64. doi:10.1111/j.1753-4887.1998.tb01732.x.

Ogden, Cynthia L., Margaret D. Carroll, Hannah G. Lawman, Cheryl D. Fryar, Deanna Kruszon-Moran, Brian K. Kit, and Katherine M. Flegal. 2016. "Trends in Obesity Prevalence Among Children and Adolescents in the United States, 1988-1994 Through 2013-2014." *JAMA* 315 (21): 2292–99. doi:10.1001/jama.2016.6361.

Pérez-Rodrigo, Carmen, Ángel Gil, Marcela González-Gross, Rosa M. Ortega, Lluis Serra-Majem, Gregorio Varela-Moreiras, and Javier Aranceta-Bartrina. 2015. "Clustering of Dietary Patterns, Lifestyles, and Overweight among Spanish Children and Adolescents in the ANIBES Study." *Nutrients* 8 (1): 11. doi:10.3390/nu8010011.

Preliminary Results

Table 2 reports item response probabilities for each indicator variable for each class in the optimal model. The largest share of students (29.6%) fall into Class 5 "High energy social butterflies". These students are unlikely to have reported daily consumption of sugary drinks and fast food. However, they are likely to report going to bed before 10:00pm, and have breakfast and meals with family members >5 times per week. These students also use >1 social media platform and own a smartphone or cellphone. Class 4 contains the second largest share of students (23.8%) and they are described as "Food conscious screen junkies". Members of Class 4 have a high probability of reporting >2 hours of television/computer use per day, they use multiple social media platforms, and own a smartphone or cellphone. While they are heavy users of electronic devices they are conscious of food and beverage choices; these students have a low probability of reporting daily consumption of sugary drinks and fast food. Class 3 (5.4%) has as similar profile to Class 4, except that these students have a higher probability of reporting consumption of foods from unhealthy retailers.

Class 2 "Role Models" (12.1%) are students who have a low probability of reporting daily sugary drink and fast food consumption. They also are likely to have family meals and breakfast >5 per week and report going to bed before 10:00pm on most nights. In addition, this class has a low probability of heavy use of electronic devices and social media.

Class 6 "Sedentary, fast food consumers" (8.6%) have a high probability of reporting multiple fast food trips per week. They also are heavy users of electronic devices and usually go to bed after 10:00pm. Class 1 (20.6%) described as "Anti-sugary drink screen junkies" do not report daily consumption of sugary drinks (and healthy drinks), but they report heavy use of electronic devices and social media.

Table 2. Proportion of total sample (n=2,791) and item response probabilities criteria for each indicator variable, conditional on latent class membership.

	Total sample (n=2,791)	Class 1 Anti-sugary drink screen junkies	Class 2 Role Models	Class 3 Food focused screen junkies	Class 4 Food conscious screen junkies	Class 5 High energy social butterflies	Class 6 Sedentary, fast food consumers
Category	Indicator variable	Item response probabilities for each indicator variable					
Sugary drinks	Soda daily	0.07	0.04	0.11	0.10	0.00	0.60
	Fruit/sport drink daily	0.21	0.28	0.40	0.50	0.25	0.56
	Energy drinks/flavored water or tea	0.36	0.59	0.62	0.65	0.70	0.48
Healthy drinks	100% juice daily	0.25	0.20	0.46	0.51	0.11	0.85
	Plain milk daily	0.06	0.07	0.17	0.25	0.01	0.66
Family cohesion	5+ family meals/week	0.43	0.77	0.78	0.85	0.87	0.56
	5+ breakfast/week	0.51	0.87	0.99	0.89	0.96	0.64
Unhealthy foods	2+ snacks at unhealthy retailers	0.39	0.14	0.98	0.31	0.13	0.82
	2+ fast food meals/week	0.89	0.40	0.82	0.86	0.50	0.95
Screen time	2+ hours/day screen time	0.92	0.07	0.76	0.95	0.91	0.90
Social media	1+ social media platforms	0.82	0.14	0.69	0.87	0.71	0.85
	Owns smartphone or cellphone	0.60	0.21	0.30	0.35	0.17	0.64
Sleep hygiene	Usually goes to bed after 10:00pm	0.36	0.10	0.92	0.17	0.05	0.81
	Class shares	20.6%	12.1%	5.4%	23.8%	29.6%	8.6%

Notes: Indicator variables with item response probabilities ≤.30 (blue) and ≥.70 (yellow) were used to describe each class.

We found that weight status was statistically associated with latent class membership $\chi^2 = 22.38$, $p < 0.001$. Regression results show statistical associations between class membership probability and sociodemographic characteristics. Gender and race were most strongly associated with latent class membership. The association between these characteristics and latent class membership probability was significantly different across the six classes.

Discussion

Our results confirm our original hypotheses and suggest that a more "person-centered" approach to the development of obesity interventions is needed. We found six distinct latent groups of students (between the ages of 10 and 13) based on their self-reported consumption of sugary drinks, healthy drinks, and unhealthy foods; family cohesion; screen time, social media use and sleep hygiene. We also found a strong association between membership to these latent groups, weight status, and sociodemographic characteristics. Each of these groups exhibits a distinct risk profile for obesity and as such different approaches for educating these groups of children about maintaining a healthy weight is essential.

Acknowledgements

The authors would like to thank the University of Connecticut and the Horizon Foundation for their support of this work.