Published in final edited form as:

J Acad Nutr Diet. 2017 May ; 117(5): 753–762. doi:10.1016/j.jand.2017.01.022.

EXPLORING THE ROLE OF SUGAR SWEETENED BEVERAGE CONSUMPTION IN OBESITY AMONG NEW YORKERS USING PROPENSITY SCORE MATCHING

Marissa Burgermaster, PhD,

Postdoctoral Research Fellow, Department of Biomedical Informatics, 622 W. 168th Street, PH 20, New York, NY 10032, At the time of the study: Program in Nutrition, Department of Health and Behavior Studies, Teachers College Columbia University, 525 West 120th Street, Box 137, New York, NY 10027

Hiershennee Bhana, MHSc,

Doctoral Candidate, Program in Nutrition, Department of Health and Behavior Studies, Teachers College Columbia University, 525 West 120th Street, Box 137, New York, NY 10027

M. Dot Fullwood, MS,

Doctoral Candidate, Health Education Program, Department of Health and Behavior Studies, Teachers College Columbia University, 525 West 120th Street, Box 114, New York, NY 10027

Diego A. Luna Bazaldua, PhD, and

Current: Postdoctoral Research Associate, Boston College, Lynch School of Education, Campion Hall, Room 332E, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, At the time of the study: Applied Statistics, Department of Human Development, Teachers College Columbia University, 525 West 120th Street, Box 118, New York, NY 10027

Elizabeth Tipton, PhD

Assistant Professor of Applied Statistics, Department of Human Development, Teachers College Columbia University, 525 West 120th Street, Box 118, New York, NY 10027

Abstract

Background—Results from clinical trials have shown that sugar-sweetened beverages (SSBs) lead to increased BMI and obesity. This relationship has yet to be explored in observational data for non-clinical populations of adults.

Objective—To compare adults who drank 4+ SSBs daily to those who drank 0 in the population of adults in New York City, and to better understand adult risk factors associated with higher daily SSB consumption and BMI.

Correspondence to: Marissa Burgermaster.

CONFLICT OF INTEREST DISCLOSURE

No authors have conflicts of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Design—Secondary analysis of cross-sectional data using propensity score matching (PSM).

Participants/Setting—2009 NYC Community Health Survey (n=9,934).

Main outcome measure—BMI

Statistical Analyses—For each of the participants who consumed 4+ SSBs daily, PSM identified matched comparisons who did not drink any SSBs. T-tests compared BMI in unadjusted and matched pairs. A post hoc analysis compared features of those likely to drink SSBs and those not.

Results—In unmatched analyses, participants who consumed 4+ SSBs daily (n=475) had higher BMI than those who consumed 0 SSBs (n=3,818; difference=1.4 kg/m², SE=0.29; t-value=4.81, p<0.001); however, when compared to similar participants using nearest neighbor with replacement matching (n=1,062), the difference between those who consumed 4+ SSBs daily and those who consumed none decreased (difference=0.37 kg/m², SE=0.36; t-value=1.01, p=0.32). Analyses also indicated that those likely to drink SSBs and those unlikely to drink SSBs differed in several important characteristics, including sex, age, race, ethnicity, SES, education, diet, and exercise.

Conclusions—The data preclude strong causal conclusions about the role of SSB in obesity. However, our results suggest that there is a subset of participants demographically and behaviorally similar with higher BMI regardless of their self-reported SSB intake. In addition to targeting SSBs, public health policies and programs should identify and address other modifiable aspects of this profile and tailor approaches to the groups identified to be most affected by high BMI.

Keywords

Sugar sweetened beverages; obesity; propensity score matching; tailoring; health disparities

Introduction

Sugar-sweetened beverages (SSBs), including non-diet soda, fruit drinks, sweetened iced teas, and sports/energy drinks, have been associated with obesity in meta-analyses and systematic reviews.^{1–3} Researchers have proposed a causal biochemical mechanism relating high sugar content to weight gain via low satiety⁴ and low compensatory adjustment within total daily caloric intake.⁵ Short-term clinical trials have found an increase in fat mass in healthy, lean adults⁶ and weight gain among obese adults⁷ with SSB supplementation. Thus, SSBs have been identified as a potential target to reverse the obesity epidemic. The American Heart Association recommends added sugars not exceed 100–150 kcal/day⁸ and identified sugar-sweetened beverages as the primary source of added sugars in the American diet.⁹ In New York City (NYC) citywide policy and media campaigns have aimed to reduce SSB consumption.¹⁰

In recent years, emphasis has been placed upon the efficient use of resources through evidence-based interventions. Although the randomized controlled trial (RCT) is considered the gold standard of evidence, it is difficult and resource-intensive to measure human dietary behavior using RCTs. Therefore, observational methods are often used to draw conclusions

to guide policy. While the generalizability of results to a broader population are difficult based on small experimental samples, selection bias due to the lack of randomization of participants to exposure or comparison limits the detection of causality in observational studies.¹¹ Statistical methods, such as covariate adjustment in multiple regression, are commonly applied to control for such confounders, but these models are limited in their ability to draw causal inferences.^{12,13}

It is important to determine if the effect of SSB on obesity seen in clinical trials is generalizable to real-life settings. However, a definitive, externally valid, cause-and-effect conclusion would require experimental trials that used random selection and a representative sample of the target population, which are not feasible to conduct. Using propensity score matching (PSM) reduces selection bias in a representative sample and, by providing a more precise description of the relationship between exposure and outcomes, it approximates causal inferences from cross-sectional data.^{12,14,15} Findings revealed in RCTs and later scaled in real world settings can better justify funding spent on large-scale health promotion programs.

PSM has been used extensively in economic, medical, and education research to estimate treatment effects with observational data.¹² PSM has not been widely applied in nutrition and diet-related research, though it has been used to assess child growth in Mexico¹⁶ and sub-Saharan Africa¹⁷ as well as the effect of school breakfast on obesity.¹⁸ However, PSM has not yet been applied to explore the relationship between SSB consumption and obesity.

Here, we explored the hypothesized link between SSB consumption and body mass index (BMI) with data from a representative sample of adult New Yorkers (residents of New York City). We first compared the influence of SSB on BMI in an unadjusted model (i.e., the simple difference between those who drink more than 4 SSBs daily and those that do not in the sample without adjusting for other differences between these groups). We then estimated the probability that each person in the population would drink 4 or more SSBs daily based on a wide range of covariates and divided the population into two groups:¹⁴ those with some probability of drinking 4 or more SSBs daily and those without. We then conducted two analyses. In Study 1, within the sample of people with some probability of drinking 4 or more SSBs daily, we compared those who actually drank 4 or more SSBs daily to those who drank none, allowing us to isolate the effect of SSBs on BMI. In Study 2, we then compared those with some probability of drinking 4 or more SSBs daily to those without, providing a better understanding of the risk factors associated with higher SSB consumption and BMI.

Methods

Data

The Community Health Survey (CHS)¹⁹ is an annual survey of NYC residents with landline or cellular telephones designed to provide self-reported data on a broad range of health topics, risk factors, and demographics from a stratified random sample of New Yorkers conducted in five languages. In 2009, data were collected from 9,934 non-institutionalized individuals over 18 years old. We used 2009 data because it was the earliest year for which SSB consumption data were collected and had the least chance to be confounded by the

NYC initiatives to regulate SSB and associated media campaigns, which began in late 2009. CHS cooperation rate for 2009 was 89.5%; most variables in CHS have <2% missing data.¹⁹ Only complete data were included in these analyses.

CHS includes BMI in m/kg^2 calculated for participants based on responses to two CHS items: “About how tall are you without shoes?” and “About how much do you weigh without shoes?” CHS includes a continuous variable representing the number of SSBs consumed daily for each participant calculated based on responses to two CHS items: “How often do you drink sugar sweetened soda? Do NOT include diet soda or seltzer” and “How often do you drink other sweetened drinks like sweetened iced tea, sports drinks, fruit punch or other fruit-flavored drinks? Do NOT include diet soda, sugar free drinks, or 100% juice.” One SSB was 12 ounces. We recoded SSB consumption as a binary variable: SSB drinkers, people who consumed 4+ SSB daily and non-SSB drinkers, people who did not consume SSBs. Experimental studies demonstrating significant effects of SSB on weight have supplemented adult ad libitum diets with 300–500 kcal of SSB, which approximately corresponds to four 12-ounce SSB.^{6,7}

The Teachers College Columbia University Institutional Review Board determined this secondary data analysis of publicly available, de-identified data to be exempt.

Propensity Score Matching

A well-implemented randomized experiment is the best method to isolate the causal effect of drinking SSBs on BMI. However, if the effect of SSBs differs across people, the causal effect estimated in an experiment might differ from that expected in the general population.^{20,21} For example, if the effect of SSB is larger for those with lower baseline BMI and an experiment only includes people with normal BMI, the effect estimated could be larger than in the general population, which has a higher average BMI.²²

An alternative approach to isolating the effect of SSBs in the population is to compare SSB drinkers in the population to non-SSB drinkers using survey data that represents a population. If the BMI for these two groups – those drinking SSBs and those not – are compared directly, however, the difference that results might not be caused by SSB consumption at all, but instead by other risk factors associated with SSBs. One approach to reduce this bias is to adjust the estimated treatment effect for differences between SSB drinkers and non-SSB drinkers on a variety of measured covariates using regression analysis. Regression approaches, however, are not adequate when the baseline differences in the two groups are large, as can easily occur in observational studies like this one.²³

An alternative approach to adjusting for measured covariates is to use PSM. In this approach, first the probabilities (or “propensity scores”) of drinking 4 or more SSBs were estimated for each person in the sample. In the current analysis, we used logistic regression to generate these propensity scores. We compared participants on 15 independent variables presented in Table 1, including: age, sex, race/ethnicity, marital status, educational attainment, household income, borough, number of young children, household size, whether the participant was born in the United States, whether the participant resided in an NYC Department of Health-designated high chronic disease neighborhood with a district public

health office (DPHO), the number of fruits and vegetables eaten the previous day, exercise habits in the past month, general health rating, and if their primary care practitioner discussed their weight with them in the past year. These independent variables represented 4 domains theorized to influence SSB consumption and BMI: demographics, nutrition, general health, and physical activity. Since the propensity scores were generated without using data on BMI or weight status, the study was designed independently of the expected outcomes; this is an important feature of propensity score analyses (as compared to regression), since it reduces researcher bias and mimics the separation of design and analysis found in experiments.²³

When all covariates that influence a participant's SSB consumption and are related to the BMI are accounted for, it is possible to get an unbiased estimate of the effect of SSB on BMI for the population. However, a difficulty with PSM, as with regression, is that the set of covariates measured in a survey may not be complete or measured well. For example, in this study, both measures of SSB consumption and measures of BMI are self-reported by participants. Furthermore, the covariate set included may omit other factors that are related to SSB consumption. In cases like this one, we can expect that all bias may not be removed, though many studies suggest that these methods may still remove most of the bias.²⁴ The PSM methods, therefore, while not as internally valid as experimental results, have higher internal validity than other more standard analyses (e.g., unadjusted or regression analyses).

After estimating the probability of drinking four or more SSBs daily, we divided the study sample into several groups, depicted in Figure 1. First, we split the sample into two larger groups: those with some probability of drinking SSBs and those without. This step is central to the analysis, for two reasons. First, it makes little sense to calculate the effect of SSB consumption on a population of people that – per the model – will not ever drink four or more SSBs per day. Second, by reducing the analysis to only those with some probability of SSB consumption, the statistical adjustments are reduced, allowing for fewer modeling assumptions and extrapolations.²⁰ Next, within the sample of those who would potentially drink 4 or more SSBs daily, we compared actual SSB-drinkers to likely SSB-drinkers using a variety of PSM estimators that further reduce differences in the measured covariates in these two groups.

In this paper, we focus on three matching procedures: sub-classification with 5 strata, nearest neighbor matching without replacement, and nearest neighbor matching with replacement. As the non-SSB drinking group was large relative to the SSB drinking group, exact matching was not possible. Nearest neighbor matching with replacement allowed cases to be matched with up to five comparisons. We assessed the degree of covariate balance in each of these methods, with the goal of finding the approach that achieved the greatest reduction in standardized mean differences (SMD) between measured covariates and average absolute SMD less than 0.10. Importantly, the outcome (BMI) was only added to the model after adequate balance on the measured covariates was achieved.

Group Comparisons

In this paper we present two studies with different analyses. First, after determining the matching method that produced the best balance in measured covariates between actual SSB

drinkers and likely SSB drinkers, we used t-tests ($\alpha=.05$) to compare average BMI between actual SSB drinkers and likely SSB drinkers. In Figure 1, this is a comparison of Groups A and B. Second, among non-SSB drinkers, we compared likely SSB drinkers and unlikely SSB drinkers (Figure 1, Groups B and C), describing differences between these groups using SMDs. Importantly, this second analysis is not possible using a regression framework, providing a third benefit for use of PSM. All analyses were conducted in R²⁵; PSM was conducted using the package ‘MatchIt’.²⁶

Results

4,293 CHS participants (43%) reported consuming either 0 or 4+ SSBs daily and had complete data for the variables included in this study. SSB drinkers (n=475) and non-SSB drinkers (n=3818) were different in sociodemographic and behavioral characteristics, summarized in Table 1. The average absolute SMD (i.e., the standardized metric used to define the extent to which groups differed on relevant covariates) was 0.291 before adjustment using matching. After propensity scores were estimated, the sample was divided: the first subset included participants consuming 4+ SSBs (Group A, n=475) and those reporting no intake of SSBs but with a similar probability of 4+ SSB consumption given their sociodemographic and behavioral covariates (Group B, n=1,062), the second subset included those neither drinking SSBs nor matched with 4+ SSB consumers (n = 2756, Group C). Nearest neighbor matching with replacement produced improved balance between Group A and B, with an average absolute SMD of 0.023; nearest neighbor matching without replacement produced an average absolute SMD between participants consuming 4+ SSBs (n=475) and matched controls (n=475) of 0.049. Sub-classification produced an average absolute SMD of 0.075 across 5 sub-strata. Although all PSM strategies yielded excellent balance between the two groups, nearest neighbor matching with replacement produced the most similar groups at the covariate level. Thus, the subsequent results and discussion will be based on the sample generated with this method. Balance statistics for nearest neighbor matching with replacement are reported in Table 1.

In Study 1, actual SSB drinkers were compared with non-SSB drinkers. The BMI difference equaled 1.4 ± 0.29 ($t=4.81$; $P<0.001$) between SSB drinkers and non-SSB drinkers when no matching method was used, but when nearest neighbor matching with replacement was used to compare the SSB drinkers and likely SSB drinkers (i.e., Group A vs. Group B), the BMI difference equaled 0.365 ± 0.36 ($t=1.01$; $P=0.32$). Differences in BMI for all models are presented in Table 2.

In Study 2, we compared the two groups of non-SSB drinkers, likely SSB drinkers and unlikely SSB drinkers on all measured covariates (i.e., groups B vs C). Analyses indicated matched (i.e., Group B) and unmatched controls (i.e., Group C) differed in their distribution of both the covariates and outcomes. This comparison is presented in Table 3. Participants in Group B had a higher average BMI and were, on average, younger, ate fewer fruits and vegetables, had larger households, and reported less physical activity compared with the unmatched control cases. Additionally, more of those in group B were male, black or Hispanic, born in the United States, never married, had a high school education, reported a household income less than 100% of the poverty line, and many lived in the Bronx. More

participants in Group C were older, white, married or widowed, had a college education, reported a household income at least 600% of the poverty line, lived in a non-DPHO neighborhood, and had no children.

Discussion

In this study, we presented a novel approach to exploring the relationship between SSBs and BMI among New Yorkers using PSM. This population is of particular interest given recent policy debate on SSB taxes in New York City.¹⁰ In unadjusted analyses, New Yorkers who reported consuming four or more SSBs daily had, on average, a higher BMI than those who consumed zero SSBs. In these analyses, we show that this difference was reduced when those who consumed four or more SSBs daily were compared via PSM only to those who were like them on a variety of measured covariates related to SSB consumption. Furthermore, likely SSB consumers were different from those unlikely to consume SSBs on a variety of specific demographic and behavioral characteristics. Together, these results suggest that SSB consumption may be a marker for a set of sociodemographic, environmental, health, and behavioral factors related to higher BMI, but that SSB might itself not be the only cause of these differences.

Our finding that SSB consumption was not related to increased BMI is inconsistent with results of meta-analyses linking SSB and obesity, such as Vartanian et al.,³ who found an overall correlation (r) of 0.08 ($p < 0.001$) for SSB consumption on body weight. However, authors of systematic reviews have noted that the differences in study design, methodologies, and data quality make it difficult to observe consistent findings.^{1,3} Existing trials suffer from many methodological limitations, including the use of self-reported dietary and physical activity data, small sample sizes, short follow-up, and lack of repeated measures of diet and lifestyle.²⁷

Our results could differ from experimental studies for a few reasons. First, the PSM approach – like regression – can only adjust for measured covariates, leaving the possibility that other variables related to SSB consumption and BMI were omitted from the analysis. Related to this is a concern with differential measurement error. It is possible participants in our study reported no SSB consumption, when in fact they do consume SSB; and/or that participants underreported their weight. However, we would anticipate similar underreporting of weight across our entire sample of overweight participants.²⁸

There are other possible explanations, however, related to the generalizability of findings from randomized experiments to a population. For example, one explanation could be a substitution effect: If those who do not consume SSB have a diet consisting of other calorically dense foods, they may have a similarly elevated BMI. Alternately, participants electing to participate in a clinical trial may be inherently different (in ways related to metabolism, motivation, or compensatory action) from those who choose not to participate. A greater proportion of the New Yorkers who were included in our actual and likely SSB drinker groups (Groups A and B) lived in DPHO neighborhoods, which are defined by higher levels of chronic disease, mortality, and health disparity, and in particular, diabetes and obesity, which are disorders of metabolism.²⁹ To disentangle some of the factors that

could have influenced results, it would be useful to examine the relationship using PSM methods on a representative longitudinal dataset that used objective measures for BMI, or in randomized trials with participants that more closely mirror the population at risk for SSB consumption.

The finding that SSB consumption is likely a marker for a constellation of factors related to elevated BMI suggests that a more holistic approach to addressing obesity may be necessary. Many health promotion efforts have focused on targeting and changing single health behaviors, such as fast food, SSB, or fruit and vegetable intake; however, this research and other studies suggest that taking a whole-diet approach may be more appropriate to address the complexities of obesity. For example, fast food consumption has been associated with poorer diet quality in overweight and obese women.³⁰ It has also been associated with a host of poor dietary patterns (e.g., lower intake of dairy, vegetables, and higher intake of SSBs) and these patterns were more strongly associated with poor dietary outcomes than fast food alone.³¹ A dietary pattern of low fruit and vegetable intake along with a high intake of meat and highly processed foods (including SSBs and fast food), has been associated with visceral adiposity³² and poor diet quality has been associated with heavily processed foods.³³ These findings suggest that SSB reduction efforts may be necessary, but not sufficient, to reduce obesity rates if the whole diet is not addressed.

Our finding that New Yorkers from lower SES and minority subgroups consumed more SSBs and had higher BMIs is consistent with research indicating racial/ethnic disparity in national obesity rates. 68% of the US population is overweight or obese;³⁴ however black and Hispanic adults have the highest age-adjusted rates.³⁵ These results are also consistent with research demonstrating that unhealthy food marketing is also disproportionately targeted at these groups.³⁶ Health promotion efforts should not only take a whole-diet approach, but also be appropriately tailored to reach these groups.

Strengths and Limitations

To our knowledge, this is the first study to use PSM to describe the relationship between SSB intake and BMI. By using this method, we reduced selection bias in an analysis of a large, representative community-based dataset. A strength of this analysis was the comparison of different matching strategies³⁷ yielding similar results; this sensitivity analysis indicates that PSM was an appropriate method for addressing our research aims. Furthermore, the excellent balance achieved with nearest neighbor matching with replacement allowed for an exploration of SSBs' relationship to BMI as well as a specific description of the group of New Yorkers who fit a sociodemographic, environmental, health, and behavioral profile associated with higher BMI.

The current study was limited by the use of self-reported weight, height, and dietary recall data in CHS. Beverages and weight have been demonstrated to be underreported.^{28,38} However, misreporting would be similar across groups and should not greatly influence comparisons. As we have discussed, a limitation of PSM is that unobserved confounding variables cannot be accounted for. In other words, although excellent balance indicated that measured covariates had very similar distributions in the Group A and Group B, it is possible that variables not collected as part of CHS also had a systematic effect on SSB

consumption. We reduced the chances of this problem by using a large number of relevant covariates across several domains in our PSM models. Finally, CHS does not capture data from households without a telephone or cellphone; nor does it capture data from adults who live in institutional housing, including college dormitories.

Conclusions

Based on our analysis of SSB consumption and BMI among New Yorkers, we propose that high SSB consumption may be an indicator of a sociodemographic, environmental, health, and behavioral profile that leads to higher BMI. While the potential for unmeasured confounders precludes a strong causal conclusion about the role of SSB in obesity, our results clearly suggest that there is a subset of New Yorkers who experience a similar set of characteristics and also have higher BMI regardless of their self-reported SSB intake. This suggests that in addition to targeting SSBs, public health programs and policies should identify and address other modifiable aspects of this profile and tailor approaches to the groups identified to be most affected by high BMI. Future research should further explore differences between experimental findings and survey findings in studies of SSB consumption, particularly attempting to understand if effects found in laboratory settings apply to those with the variety of complex risk factors found in diverse populations.

Acknowledgments

FUNDING/SUPPORT DISCLOSURE

Corresponding author was supported by a National Institutes of Health T32 training grant (NHLBI HL007343) and is currently supported by a National Institutes of Health T15 training grant (NLM LM007079).

References

1. Malik VS, Popkin BM, Bray GA, Després J-P, Hu FB. Sugar-sweetened beverages, obesity, type 2 diabetes mellitus, and cardiovascular disease risk. *Circulation*. 2010; 121(11):1356–1364. [PubMed: 20308626]
2. Te Morenga L, Mallard S, Mann J. Dietary sugars and body weight: systematic review and meta-analyses of randomised controlled trials and cohort studies. *BMJ*. 2013; 346:e7492.
3. Vartanian LR, Schwartz MB, Brownell KD. Effects of soft drink consumption on nutrition and health: a systematic review and meta-analysis. *Am J Public Health*. 2007; 97(4):667–675. [PubMed: 17329656]
4. DiMeglio DP, Mattes RD. Liquid versus solid carbohydrate: effects on food intake and body weight. *Int J Obes*. 2000; 24(6):794–800.
5. Mattes RD. Dietary compensation by humans for supplemental energy provided as ethanol or carbohydrate in fluids. *Physiol Behav*. 1996; 59(1):179–187. [PubMed: 8848479]
6. Sartor F, Jackson MJ, Squillace C, et al. Adaptive metabolic response to 4 weeks of sugar-sweetened beverage consumption in healthy, lightly active individuals and chronic high glucose availability in primary human myotubes. *Eur J Nutr*. 2013; 52(3):937–948. [PubMed: 22733000]
7. Raben A, Vasilaras TH, Møller AC, Astrup A. Sucrose compared with artificial sweeteners: different effects on ad libitum food intake and body weight after 10 wk of supplementation in overweight subjects. *Am J Clin Nutr*. 2002; 76(4):721–729. [PubMed: 12324283]
8. Johnson RK, Appel LJ, Brands M, et al. Dietary sugars intake and cardiovascular health a scientific statement from the American Heart Association. *Circulation*. 2009; 120(11):1011–1020. [PubMed: 19704096]
9. Block G. Foods contributing to energy intake in the US: data from NHANES III and NHANES 1999–2000. *J Food Compos Anal*. 2004; 17(3):439–447.

10. Alberti, P., Noyes, P. Sugary Drinks: How Much Do We Consume?. New York City Department of Health and Mental Hygiene; 2011. <http://www1.nyc.gov/assets/doh/downloads/pdf/dpho/dpho-sugary-drinks-report.pdf> [Accessed December 5, 2016]
11. Shadish, WR., Cook, TD., Campbell, DT. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin; 2001.
12. d'Agostino RB. Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*. 1998; 17(19):2265–2281. [PubMed: 9802183]
13. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol*. 1974; 66(5):688.
14. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983; 70(1):41–55.
15. Stuart EA, Cole SR, Bradshaw CP, Leaf PJ. The use of propensity scores to assess the generalizability of results from randomized trials. *J R Stat Soc Ser A Stat Soc*. 2001; 174(2):369–386.
16. Leroy JL, García-Guerra A, García R, Dominguez C, Rivera J, Neufeld LM. The Oportunidades program increases the linear growth of children enrolled at young ages in urban Mexico. *J Nutr*. 2008; 138(4):793–798. [PubMed: 18356337]
17. Remans R, Pronyk PM, Fanzo JC, et al. Multisector intervention to accelerate reductions in child stunting: an observational study from 9 sub-Saharan African countries. *The Am J Clin Nutr*. 2011; 94(6):1632–1642. [PubMed: 22030229]
18. Sudharsanan N, Romano S, Cunningham SA. School Breakfast Receipt and Obesity among American Fifth-and Eighth-Graders. *J Acad Nutr Diet*. 2016; 116(4):599–607. [PubMed: 27046061]
19. New York City Department of Health and Mental Hygiene. Community Health Survey 2009; public use dataset. [accessed on October 12, 2015]
20. Ho DE, Imai K, King G, Stuart EA. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal*. 2007; 15(3):199–236.
21. Tipton E. Improving Generalizations From Experiments Using Propensity Score Subclassification Assumptions, Properties, and Contexts. *J Educ Behav Stat*. 2012 1076998612441947.
22. Flegal KM, Troiano RP. Changes in the distribution of body mass index of adults and children in the US population. *Int J Obes*. 2000; 24(7):807–818.
23. Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med*. 2007; 26(1):20–36. [PubMed: 17072897]
24. Steiner PM, Cook TD, Shadish WR, Clark MH. The importance of covariate selection in controlling for selection bias in observational studies. *Psychol Methods*. 2010; 15(3):250. [PubMed: 20822251]
25. Team RC. [Accessed October 12, 2015] R: A language and environment for statistical computing. 2012. <http://www.r-project.org>
26. Ho DE, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference (version 2.211)[software]. *J Stat Softw*. 2009
27. Hu FB. Resolved: there is sufficient scientific evidence that decreasing sugar-sweetened beverage consumption will reduce the prevalence of obesity and obesity-related diseases. *Obes Rev*. 2013; 14(8):606–619. [PubMed: 23763695]
28. Gorber SC, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev*. 2007; 8(4):307–326. [PubMed: 17578381]
29. Singla P, Bardoloi A, Parkash AA. Metabolic effects of obesity: A review. *World J Diabetes*. 2010; 1(3):76–88. [PubMed: 21537431]
30. Wilcox S, Sharpe PA, Turner-McGrievy G, Granner M, Baruth M. Frequency of consumption at fast-food restaurants is associated with dietary intake in overweight and obese women recruited from financially disadvantaged neighborhoods. *Nutr Res*. 2013; 33(8):636–646. [PubMed: 23890353]

31. Poti JM, Duffey KJ, Popkin BM. The association of fast food consumption with poor dietary outcomes and obesity among children: is it the fast food or the remainder of the diet? *Am J Clin Nutr.* 2014; 99(1):162–171. [PubMed: 24153348]
32. Vilela AA, Sichieri R, Pereira RA, et al. Dietary patterns associated with anthropometric indicators of abdominal fat in adults. *Cad Saude Publica.* 2014; 30(3):502–510. [PubMed: 24714940]
33. Moubarac JC, Martins AP, Claro RM, Levy RB, Cannon G, Monteiro CA. Consumption of ultra-processed foods and likely impact on human health. Evidence from Canada. *Public Health Nutr.* 2013; 16(12):2240–2248. [PubMed: 23171687]
34. Ogden CL, Carroll MD, Fryar CD, Flegal KM. Prevalence of Obesity Among Adults and Youth: United States, 2011–2014. *NCHS Data Brief.* 2015; (219):1–8.
35. CDC Health Disparities & Inequalities Report – United States, 2011. *MMWR Suppl.* 2011; 60:1–116.
36. Harris, J. Food and beverage marketing targeted to black and Latino youth and families. Paper presented at: 2015 National Conference on Health Communication, Marketing, and Media; August 11–13, 2015; Atlanta, GA. <https://cdc.confex.com/cdc/nphic15/webprogram/Paper36189.html>
37. Mayne SL, Lee BK, Auchincloss AH. Evaluating Propensity Score Methods in a Quasi-Experimental Study of the Impact of Menu-Labeling. *PLoS One.* 2015; 10(12):e0144962. [PubMed: 26677849]
38. Gemming L, Mhurchu CN. Dietary under-reporting: what foods and which meals are typically under-reported? *Eur J Clin Nutr.* 2016; 70(5):640–641. [PubMed: 26669571]

Figure 1. Study Subgroups of New York City Community Health Survey Participants by SSB Consumption

Notes: SSBs, Sugar-sweetened beverages. Groups based on nearest neighbor matching with replacement (ratio=5).

Table 1

Characteristics of SSB Drinkers (≥ 4 SSBs/day) and Non-SSB Drinkers in New York City Before Matching and Balance on Participant Characteristics Before and After Propensity Score Matching

	Means SSB Drinkers (Mean±SD; n=475)	Means Non-SSB Drinkers (Mean±SD; n=3818)	Balance before matching	Balance after matching ^a
Age in years	41.42±15.40	57.07±15.90	-1.02	-0.02
FV servings yesterday	1.60±1.46	2.87±2.55	-0.87	0.00
Adults in household	2.69±1.62	2.14±1.44	0.34	-0.02
Race				
White	0.27±0.44	0.56±0.50	-0.66	-0.03
Black	0.36±0.48	0.16±0.36	0.42	0.02
Hispanic	0.32±0.47	0.20±0.40	0.27	0.01
Asian/Pacific Islander	0.03±0.18	0.07±0.26	-0.21	-0.01
Other	0.01±0.12	0.01±0.12	0.01	-0.01
Self-reported general health ^b	2.73±1.08	2.70±1.22	0.03	-0.01
PCP discussed weight this year ^c	0.38±0.49	0.37±0.48	0.03	0.00
Physical activity this month ^c	0.61±0.49	0.74±0.44	-0.26	0.00
Male gender ^c	0.48±0.50	0.36±0.48	0.24	-0.07
Born in United States ^c	0.75±0.44	0.63±0.48	0.26	0.03
Marital Status				
Married	0.26±0.44	0.39±0.49	-0.30	-0.03
Divorced	0.12±0.32	0.16±0.36	-0.13	0.01
Widowed	0.05±0.22	0.15±0.36	-0.46	0.01
Separated	0.07±0.26	0.05±0.21	0.10	-0.03
Never married	0.42±0.49	0.21±0.41	0.42	0.01
Part of unmarried couple	0.08±0.26	0.04±0.19	0.15	0.04
Education				
Less than high school	0.17±0.38	0.14±0.34	0.09	0.04
High school graduate	0.33±0.47	0.20±0.40	0.28	-0.04
Some college	0.27±0.44	0.19±0.39	0.17	-0.01
College grad	0.23±0.42	0.48±0.50	-0.57	0.01
Household poverty				
<100%	0.26±0.44	0.17±0.37	0.21	-0.06
100–200%	0.21±0.41	0.16±0.36	0.14	0.02
200–<400%	0.16±0.37	0.15±0.35	0.04	-0.01
400–<600%	0.13±0.33	0.16±0.37	-0.09	0.02
>600%	0.13±0.34	0.28±0.45	-0.44	0.06
Don't know	0.11±0.31	0.09±0.29	0.06	-0.02
Borough				

	Means SSB Drinkers (Mean±SD; n=475)	Means Non-SSB Drinkers (Mean±SD; n=3818)	Balance before matching	Balance after matching ^a
Bronx	0.24±0.43	0.14±0.34	0.25	0.04
Brooklyn	0.29±0.45	0.25±0.44	0.07	0.04
Manhattan	0.18±0.38	0.28±0.45	-0.26	-0.06
Queens	0.20±0.40	0.24±0.43	-0.09	-0.03
Staten Island	0.08±0.28	0.09±0.28	-0.02	0.00
DPHO Neighborhood				
South Bronx	0.10±0.30	0.05±0.21	0.19	0.01
East/Central Harlem	0.07±0.25	0.03±0.17	0.14	-0.04
North/Central Brooklyn	0.12±0.32	0.06±0.23	0.19	0.06
All other neighborhoods	0.72±0.45	0.87±0.34	-0.34	-0.02
Children under 6 in household				
0	0.79±0.41	0.90±0.30	-0.27	-0.02
1	0.13±0.34	0.07±0.26	0.18	0.02
2	0.06±0.24	0.03±0.16	0.14	0.01
3	0.02±0.13	0.00±0.06	0.10	-0.02
>3	0.00±0.05	0.00±0.03	0.03	0.01
Propensity Score				0.01
Average Absolute Balance			0.24	0.02

Notes: SSBs, Sugar-sweetened beverages; SD, standard deviation; FV, fruit and vegetable; PCP, primary care practitioner; DPHO, district public health office – designated high health need area;

^aBalance after nearest neighbor matching with replacement (ratio=5),

^bHealth status self-reported on a scale of 1–5,

^cDichotomous variable.

Table 2

Differences between Actual and Potential SSB Drinkers from the NYC Community Health Survey in Unadjusted and Propensity Score Matched Comparisons

Type of Matching	BMI (m/kg ²)				
	n _S	n _C	SSB Drinkers (Mean±SD)	Comparisons (Mean±SD)	SE
Unadjusted	475	3818	28.10±6.51	26.69±5.94	1.40
Nearest Neighbor					0.29 ***
With replacement	475	1062	28.10±6.51	27.73±6.61	0.37
Without replacement	475	475	28.10±6.51	27.97±6.66	0.13
Subclassification					0.43
Stratum 1 (0 < p < 0.012)	12	847	25.69±4.28	25.16±4.73	0.53
Stratum 2 (0.012 < p < 0.030)	13	845	25.91±5.50	26.07±5.26	-0.16
Stratum 3 (0.030 < p < 0.065)	40	819	28.78±5.53	26.97±6.01	1.81
Stratum 4 (0.065 < p < 0.168)	87	771	28.25±6.82	27.54±6.31	0.71
Stratum 5 (0.168 < p < 1)	323	536	28.15±6.64	28.46±7.16	-0.31

Notes: SSB, Sugar-sweetened beverage; S, actual SSB drinkers; C, Comparisons; SD, standard deviation; BMI, body mass index; Unadjusted analyses reflect the simple difference between those that reported drinking 4 SSBs/day and those that do not in the sample without adjusting for other differences between these groups; Strata were created such that each contained 1/5th of the total sample (ranges of estimated probabilities of drinking ≥4 SSBs for each stratum provided in parentheses);

p<0.001

Table 3

Comparison of Two Samples of Non-SSB Drinkers from NYC Community Health Survey on Relevant Covariates

	Means Unlike SSB Drinkers ^a (Mean ±SD; n=2756)	Means Like SSB Drinkers ^b (Mean ±SD; n=1062)	Comparison (SMD ^c)
Age in years	60.21±14.82	48.94±15.74	11.27
FV servings yesterday	3.20±2.78	2.03±1.52	1.19
Adults in household	2.03±1.37	2.43±1.59	-0.40
Race			
White	0.63±0.48	0.39±0.49	0.24
Black	0.12±0.32	0.26±0.44	-0.14
Hispanic	0.17±0.37	0.28±0.45	-0.12
Asian/Pacific Islander	0.08±0.27	0.05±0.21	0.03
Other	0.01±0.11	0.02±0.14	-0.00
Self-reported general health	2.68±1.22	2.75±1.23	-0.07
PCP discussed weight this year	0.36±0.48	0.39±0.49	-0.02
Physical activity this month	0.77±0.42	0.65±0.48	0.13
Female	0.67±0.47	0.58±0.49	0.09
Born in United States	0.61±0.49	0.70±0.46	-0.09
Marital Status			
Married	0.42±0.49	0.33±0.47	0.09
Divorced	0.16±0.36	0.15±0.36	0.01
Widowed	0.18±0.38	0.09±0.28	0.09
Separated	0.04±0.19	0.07±0.26	-0.04
Never married	0.18±0.39	0.30±0.46	-0.12
Part of unmarried couple	0.03±0.16	0.06±0.24	-0.04
Education			
Less than high school	0.13±0.33	0.17±0.37	-0.04
High school graduate	0.18±0.38	0.26±0.44	-0.08
Some college	0.17±0.38	0.24±0.43	-0.07
College grad	0.53±0.50	0.33±0.47	0.20
Household poverty			
<100%	0.15±0.35	0.22±0.42	-0.08
100–200%	0.15±0.35	0.19±0.39	-0.04
200–<400%	0.14±0.35	0.17±0.37	-0.03
400–<600%	0.16±0.37	0.15±0.35	0.02
>600%	0.32±0.47	0.18±0.39	0.14
Don't know	0.09±0.29	0.10±0.29	-0.01
Borough			
Bronx	0.11±0.32	0.19±0.40	-0.08
Brooklyn	0.25±0.43	0.27±0.44	-0.02
Manhattan	0.30±0.46	0.23±0.42	0.07

	Means Unlike SSB Drinkers ^a (Mean ±SD; n=2756)	Means Like SSB Drinkers ^b (Mean ±SD; n=1062)	Comparison (SMD ^c)
Queens	0.25±0.43	0.23±0.42	0.02
Staten Island	0.09±0.29	0.08±0.27	0.01
DPHO Neighborhood			
South Bronx	0.03±0.18	0.08±0.27	−0.05
East/Central Harlem	0.02±0.15	0.06±0.23	−0.03
North/Central Brooklyn	0.04±0.21	0.09±0.28	−0.04
All other neighborhoods	0.90±0.30	0.78±0.42	0.12
Children under 6 in household			
0	0.93±0.26	0.84±0.37	0.09
1	0.05±0.23	0.11±0.32	−0.06
2	0.02±0.14	0.04±0.19	−0.02
3	0.00±0.04	0.01±0.09	−0.00
3	0.00±0.02	0.00±0.04	−0.00
BMI (m/kg²)	26.30±5.61	27.73±6.61	−1.43

Notes: SSB, sugar-sweetened beverage; SD, standard deviation; SMD, standard mean difference; FV, fruit and vegetable; PCP, primary care practitioner; DPHO, district public health office; BMI, body mass index;

^aUnlike SSB drinkers were not matched using nearest neighbor matching with replacement,

^bLike SSB drinkers were matched using nearest neighbor matching with replacement,

^cBold indicates SMD > |0.08|.