A Risk Analysis Model of the Relationship Between Beverage Consumption from School Vending Machines and Risk of Adolescent Overweight

Richard A. Forshee,1* Maureen L. Storey,1 and Michael E. Ginevan2

Risk analysis is a widely used tool to understand problems in food safety policy, but it is seldom applied to nutrition policy. We propose that risk analysis be applied more often to inform debates on nutrition policy, and we conduct a risk assessment of the relationship of regular carbonated soft drink (RCSD) consumption in schools and body mass index (BMI) as a case study. Data for RCSD consumption in schools were drawn from three data sets: the Continuing Survey of Food Intake by Individuals 1994–1996, 1998 (CSFII), the National Health and Nutrition Examination Survey 1999–2000 (NHANES), and the National Family Opinion (NFO) WorldGroup Share of Intake Panel (SIP) study. We used the largest relationship between RCSD and BMI that was published by prospective observational studies to characterize the maximum plausible relationship in our study. Consumption of RCSD in schools was low in all three data sets, ranging from 15 g/day in NFO-SIP to 60 g/day in NHANES. There was no relationship between RCSD consumption from all sources and BMI in either the CSFII or the NHANES data. The risk assessment showed no impact on BMI by removing RCSD consumption in school. These findings suggest that focusing adolescent overweight prevention programs on RCSD in schools will not have a significant impact on BMI.

KEY WORDS: Adolescent; body mass index; carbonated beverages; food vending machine; obesity

1. INTRODUCTION

Risk analysis is a widely used tool to understand problems in food safety policy, but it is seldom applied to nutrition policy. Food safety decisions are mandated to be “risk-based” and many national and international authoritative bodies have published extensive guidance on the appropriate use of risk analysis for food safety questions. The major nutrition policy institutions, such as the Dietary Guidelines Advisory Committee, the World Health Organization, and the Food and Agriculture Organization of the United Nations, have no similar mandate to follow a risk-based or analogous approach for nutrition guidance. Major policy documents are published without any analysis linking the exposure and hazard (or benefit) assessments to create an appropriate risk characterization. It is even rarer to see a formal discussion of variability and uncertainty, risk-risk analysis of competing proposals, or cost-benefit analysis.

This article is intended to open a discussion on the appropriate use of risk analysis in nutrition policy. As an example, we apply the National Academies of Science (NAS) four-step risk assessment process(1) to carry out a risk assessment for regular carbonated soft drink (RCSD) consumption in schools and body mass index (BMI) and prevalence of overweight. The four steps are:

1 Center for Food, Nutrition, and Agriculture Policy, University of Maryland—College Park, MD, USA.
2 Exponent, Inc., Washington, DC, USA.
* Address correspondence to Richard A. Forshee, Center for Food, Nutrition, and Agriculture Policy, University of Maryland—College Park, 0220, Symons Hall, College Park, MD 20742; rforshee@umd.edu.
1. **Hazard identification** evaluates whether exposure to an agent (RCSD consumption from school vending machines) causes an increase in some adverse health condition of interest (BMI or risk of overweight).

2. **Dose-response assessment** determines the form of the relationship between increasing dose of the agent identified in Step 1 and the incidence of the adverse health effect of interest. We use peer-reviewed estimates of the association between RCSD consumption and BMI to define the dose response in our risk assessment.

3. **Exposure assessment** estimates the exposure increment or decrement that would be expected as a result of some activity or process, in our case the removal of vending machines from schools. According to the NAS: “Exposure assessment is often used to identify feasible prospective control options and to predict effects of available control technologies on exposure.” We use consumption data from two large, nationally representative food consumption surveys conducted by the U.S. government to conduct our exposure assessment.

4. **Risk characterization** is the process of estimating the incidence of a health effect under various conditions of human exposure described in the exposure assessment. We use a deterministic approach to link the dose-response assessment and the exposure assessment to estimate the potential changes in the distribution of BMI and prevalence of overweight.

We present a novel expansion and use of the NAS risk analysis paradigm to assess issues of nutrition and their potential impact on public health. We believe the risk analysis paradigm could be widely applied to issues in nutrition and public health.

Adolescent overweight is a serious and growing problem in the United States. The percentage of adolescents classified as overweight—defined as at or above the 95th percentile of BMI for age—"increased 5 percentage points among 12- through 19-year-olds from 10.5% to 15.5% between NHANES III [1988–1994] and NHANES 1999–2000."(2) The increase was most striking for non-Hispanic black and Mexican-American adolescents. There was a 10 percentage point increase in overweight prevalence for both of these groups. The change for non-Hispanic white adolescents was not statistically significant in this time period.(2)

Overweight adolescents are at higher risk of medical conditions such as hyperlipidemia, glucose intolerance, hypertension, and sleep apnea.(3) Overweight adolescents are more likely to be overweight as adults, and they are at higher risk of cardiovascular disease and diabetes. All-cause mortality is higher among men who were obese during adolescence.(4)

The cause or causes of the increased rate of adolescent overweight is in dispute. In general, average total energy consumption for adolescents has not changed from the early 1970s to 2000, except for an increase among adolescent females.(5,6) Meanwhile, the percentage of students in grades 9–12 enrolled in daily physical education classes fell from 41.6% in 1991 to 32.2% in 2001.(7) Some specific foods have also been indicted as potential causes of the increase in BMI. For example, there is considerable controversy over carbonated soft drink consumption and the presence of certain snacks and soft drinks in school vending machines.

Five recent prospective studies have examined the relationship between consumption of RCSDs and BMI among children and adolescents. The studies show inconsistent results, and the relationship between sugar-sweetened beverages and BMI ranges from not statistically significant to a weak relationship.(8–12)

The scientific literature on the relationship between school vending machines and health is limited. French and colleagues have examined several interventions that have shown that pricing policies can affect the foods purchased from vending machines,(13) but they did not test for direct effects on health outcomes. A Medical Subject Headings (MeSH) search for “food vending machine” and “obesity” identified four articles published between 1994 and November 2004. One was a recent study conducted in the United Kingdom that found no “link between consumption of confectionery purchased from vending machines and ‘poor’ dietary practice or ‘undesirable’ lifestyle habits.”(14) The others were a policy statement,(15) a policy brief,(16) and a randomized controlled trial examining whether high fasting plasma ghrelin concentrations (FxGhr) might predict high ad libitum food intake.(17)

Despite the limited scientific database, professional organizations, state legislatures, and local school boards are implementing or considering bans or limitations on school vending machines containing “unhealthy” snacks, particularly RCSD. For example, the American Academy of Pediatrics Committee on School Health Policy Statement on Soft Drinks in
Schools states, “[a] clearly defined, district-wide policy that restricts the sale of soft drinks will safeguard against health problems as a result of overconsumption.”\(^{(18\text{, p.}152)}\)

The objectives of this study are to (1) provide estimates from multiple data sources to establish the plausible range of soft drink consumption from school vending machines and (2) develop a risk assessment model of RCSD from school vending machines and BMI. These objectives define the scope of the risk assessment.

2. DATABASES AND METHODS

To characterize the risk of increased BMI and adolescent overweight associated with consumption of RCSD, we used the consumption data from two large, nationally representative food consumption surveys conducted by the U.S. government. Evaluation of consumption of RCSD in schools is the exposure assessment in the NAS paradigm. The hazard characterization relationship between RCSD consumption and BMI is drawn from the published scientific literature using prospective observational studies. We used the maximum interindividual variability in consumption of RCSD and the largest plausible published relationship between RCSD consumption and BMI in an exposure distribution and hazard characterization to estimate the upper plausible impact of RCSD consumption on BMI.

2.1. Consumption Data

Data for the analysis were from the Continuing Survey of Food Intake by Individuals 1994–1996, 1998 (CSFII), the National Health and Nutrition Examination Survey 1999–2000 (NHANES), and the National Family Opinion (NFO) WorldGroup Share of Intake Panel (SIP) study. CSFII and NHANES represented all noninstitutionalized persons over two years of age residing in the United States during their respective time periods. CSFII collected dietary information based on two 24-hour, multipass dietary recall (24HR) instruments administered on nonconsecutive days. We used the 2-day averages for all of the analyses in this study. Detailed information about the methods used to collect the CSFII data has been published elsewhere.\(^{(19)}\) NHANES collected dietary information on a single multipass 24HR instrument.\(^{(20)}\) We restricted our analysis to respondents aged 13–18 years—the typical age range for middle and high school students in the United States. CSFII contained 536 male respondents and 549 female respondents aged 13–18 years, and NHANES contained 839 male respondents and 824 female respondents aged 13–18 years.

Using the CSFII Individual Food File, we coded variables for the source (i.e., vending machine, store, restaurant, etc.) of each beverage and created binary identifier variables based on the beverage codes from the CSFII documentation to indicate whether a particular food item fell into one of four beverage categories (regular fruit drink/ade, diet fruit drink/ade, RCSDs, and diet carbonated soft drinks (DCSD)). We coded a binary variable to identify food items that were consumed away from home during school hours (defined as 7 a.m. to 4 p.m. for the purposes of this analysis). Both weekdays and weekends were included. Using the binary identifier variables for beverage and source, we identified beverage consumption by source and aggregated the data to the individual level for Day 1 and Day 2. Beverages that had a source of vending machine and that were consumed away from home during school hours were the operational definition of beverages from school vending machines. That file was then aggregated to a 2-day mean consumption file, and data on each respondent’s height, weight, and BMI were merged from the CSFII personal demographics file.

The same procedure was applied to the NHANES data with one exception. NHANES did not include a source of food variable, but it did include a variable identifying the place the food was eaten. We used the “place eaten” variable to identify beverages that were consumed at school and at home. The “consumed at school” variable is not a perfect estimate of beverages purchased from vending machines. It includes beverages purchased elsewhere and carried to school and it excludes beverages purchased from school vending machines but consumed elsewhere. However, it does provide another independent estimate of the amount of beverages consumed by students while at school. Data from NHANES 2001–2002 were not included because the publicly released data did not include the “place eaten” variable.

CSFII and NHANES have been widely used to monitor food consumption as data sources for risk assessments, particularly for food safety. The surveys provide guidance for nutrition policies and programs. Both surveys, but particularly NHANES, are also designed to allow the estimation of associations between diet and health.\(^{(21\text{,}22)}\) For these reasons, we believe the data are appropriate and adequate to perform the risk analysis we present.
A third independent data set used food diaries to examine consumption of RCSD from school vending machines. The NFO-SIP is a mail sample of 12,000 persons per year. Participants keep 2-week diaries of all beverages consumed, excluding tap water, and the location where these were consumed. In a survey conducted during the 2001–2002 school year, consumption of RCSD purchased from secondary school vending machines were estimated for a demographically balanced sample of 2,716 students aged 12–18 years. Two-week consumption diaries were kept by students for the school year (September 2001 through May 2002). Seasonal shifts in beverage consumption were reflected by the data and are similar to NFO data from previous surveys.(23)

Descriptive statistics were calculated in Stata 8 using the “svymean” procedure and the 2-day sample weight in CSFII or the 2-year Mobile Examination Center weight in NHANES. The distributions of consumption were calculated using “centile” and “cumul” (cumulative distribution) procedures.

A simple risk analysis model was developed to determine the upper plausible impact that eliminating RCSD consumption from school vending machines could have on BMI and overweight in the adolescent population. Using the CSFII and NHANES data on RCSD consumption and measured BMI as a baseline, hypothetical BMI values for all respondents were calculated by excluding RCSD from school vending machines (CSFII) or RCSD consumed at school (NHANES).

2.2. Relating RCSD Consumption to BMI

Risk characterization estimates the incidence of a health effect under various conditions of human exposure described in the exposure assessment. It is a way of combining information from the dose-response estimation and the exposure assessment to estimate potential changes in risk.

The relationship between RCSD consumption and BMI was based on the largest estimate from five recent prospective observational studies. Among adolescents, the estimated relationships ranged from a maximum of 0.24 kg/m² per serving/day (n = 548, p = 0.03) to a minimum of 0.02 kg/m² per serving/day (n > 10,000, p = 0.096). It would be reasonable to reflect the uncertainty of the relationship by choosing a “most likely” value in the middle of the range of estimates or by using a probability distribution reflecting the range of estimates in a Monte Carlo simulation. However, we have chosen to use the largest estimate in order to estimate the upper plausible impact of RCSD consumption on BMI. This does not suggest that the 0.24 kg/m² per serving/day estimate is the best or most likely estimate, but suggests only that it is the largest. In a sensitivity test, we use the lower estimate to bound the range of changes in BMI.

If there is a relationship between RCSD and BMI, it should follow a dose-response curve of some shape. That is, the more RCSD consumed, the greater the impact on BMI. This is the implicit assumption in the epidemiological prospective observational studies we reviewed. The estimated associations between RCSD and BMI reported in the literature are usually estimated as a linear relationship, and we have followed that convention in our article. The studies we relied on for the dose-response estimation in the base model and one sensitivity test(8,9) controlled for a wide range of other factors including total energy in the diet and physical activity. We assumed that the models were properly specified and that the reported associations were unbiased estimates. Given these assumptions, the prospective observational studies provided estimates of the predicted change in BMI per unit serving change (increase or decrease) in RCSD consumption between two time periods. Notice that these models used the change in RCSD consumption between two time periods. Once RCSD consumption from a source has been reduced to zero, there can be no further reduction of RCSD consumption—and consequently no further reduction in BMI—in future time periods. This eliminates the need for us to model an iterative process, since we are modeling the maximum reduction.

The intervention we are considering would eliminate school vending machines as a source of RCSD and so would presumably reduce RCSD consumption by the amount currently consumed from that source. It is possible that the reduction would be less if students substituted RCSD from other sources, but we have assumed that no substitution occurs. Consequently, to calculate the new hypothetical BMI, we multiplied the dose-response estimation from the epidemiological studies by the exposure estimation from CSFII or NHANES and subtracted that product from the original BMI (1).

\[
\text{Hypothetical BMI} = \text{Original BMI} - 0.24 \times \text{RCSD}_{\text{School}} \quad (1)
\]

Since the epidemiological models controlled for other factors, we do not need to include those in our model. Nothing about the form of Equation (1)
guarantees a small reduction in BMI. The reduction is dependent on the dose-response estimation and the exposure estimation. If the estimated dose response was large or if the exposure assessment was large, the equation could lead to substantial changes in BMI.

This Equation (1) links our dose-response estimate from peer-reviewed prospective, observational studies with our exposure estimate from nationally representative food consumption surveys. A serving of sugar-sweetened beverages did not have a precise quantitative definition in the Ludwig et al. study because the dietary instrument was semi-quantitative. For the purposes of our base analysis, we defined a serving as 12 oz (370 g) and scaled RCSD consumption from school accordingly. In a sensitivity test, we used an 8-oz serving size, which effectively increases exposure by one-third for consumers.

Cumulative density function graphs of the original BMI distribution and the hypothetical distribution resulting from the elimination of RCSD school consumption were produced using the “cumul” procedure in Stata. Tables of the distribution of the original and hypothetical BMI values at every 10th percentile were produced using the “centile” procedure. The percentage of the sample that was overweight or underweight was calculated by comparing the original BMI and the hypothetical BMI values to the Centers for Disease Control and Prevention (CDC) age gender-specific cutoff points from the BMI-for-age growth charts. As recommended by the CDC, the 85th percentile was used as the cutoff point for at risk overweight and the 5th percentile was used as the cutoff point for underweight.

2.3. Variability and Uncertainty

Any quantitative evaluation of risk needs to consider the variability and uncertainty of model inputs. Dietary intake data, including data for RCSD consumption, have considerable variability and, at the individual level, significant uncertainty as well. 24HR data, such as that in CSFII and NHANES, have been shown to overestimate consumption in the higher percentiles compared to long-term consumption patterns. Moreover, intakes are correlated within individuals. That is, if a person has zero RCSD intake for a given day, it is more likely that, if subsequent days were observed, the intake on these days would also be zero or low. Similarly, if a person has a very high RCSD intake for a given day, it is more likely that intake on subsequent days would also be high. In addition, dietary intakes for different items are correlated within individuals. That is, persons with high intakes of DCSD will generally have much lower intakes for RCSD.

To address these complexities, we take two approaches in our analyses. First, we model the impact of eliminating RCSD consumed at school by assuming that the 2-day average intake (for CSFII) or the single 24HR (for NHANES) for each individual accurately reflects the long-term average intake for that individual. This assumption results in an artificially wide range of intakes with higher consumption in the upper percentiles because it assumes that all observed variation in intakes reflects interindividual variation, when in fact it reflects both interindividual variability and intra-individual variability.

Our second approach is to employ a Monte Carlo bootstrap approach. In our bootstrap simulation study, we randomly selected a single record to provide a set of personal characteristics and an initial set of intakes for all beverage categories. We then randomly selected, with replacement, an additional 364 sets of intakes from other individuals in the sample. The 365 randomly selected days were averaged to estimate long-term average intakes for a simulated individual. We recognize that this is longer than the typical school year of about 180 days, but as with other assumptions in this model we chose the maximum plausible consumption values. The process was repeated 10,000 times to create 10,000 simulated individuals. This, in effect, assumed that there was no interindividual variability and that all observed variation was a result of a common interindividual variability. That is, it assumed that variation across individuals is equivalent to variation within an individual. While clearly incorrect, it gives some idea of a lower bound on interindividual variation in long-term overall intake.

We also note that there are some significant uncertainties in our modeling. The largest is arguably the form of Equation (1), which was used to calculate the impact on BMI of eliminating RCSD consumption from school vending machines. There was a fairly broad range of equations that could be assumed when making such estimates, and several others are mentioned in our discussion. However, the one used here as our Equation (1) was very much an upper-end assumption in that it yielded a relatively large reduction in BMI for a given reduction in RCSD intake. Since even this upper-end assumption yields minimal impact on BMI, we did not attempt to address the issue of model uncertainty and instead suggest that, while effects may in fact be even smaller than those calculated here, they are not likely to be much larger.
2.4. Limitations

Each of the estimates of RCSD consumption from school vending machines has some flaws. The CSFII measure of RCSD purchased from vending machines may include purchases that were not made at school. The NHANES measure of RCSD consumed at school may include RCSD purchased from sources other than vending machines (stores, snack bars, fundraisers, etc.) or brought from home. It also may not include RCSD purchased from school vending machines but consumed elsewhere. We attempted to address these limitations by using each estimate in separate models in an attempt to capture the most likely range of possible consumption values.

The 24HR does not reflect long-term intake that may lead to some attenuation in the estimate of diet and health relationships. Also, because diet is self-reported, there may be some underreporting and the underreporting may be systematically related to certain personal characteristics. However, the 5-pass method is designed to minimize underreporting. We addressed some of the potential underreporting in our sensitivity tests, which assumed a smaller serving size and, consequently, a larger number of servings. We addressed the long-term intake limitation by estimating models using both the original data, which has greater variability than one would expect from long-term intake, and a bootstrap simulation that used random draws with replacement from the data to construct simulated intake over the course of a year.

3. RESULTS

3.1. RCSD Consumption

Average per capita consumption of RCSD purchased from school vending machines was 2.5 oz/week (77 g/week) or 0.5 oz/day (15 g/day) in NFO-SIP. Among the 20% of all students who drank beverages purchased from vending machines, the intake averaged 12.5 oz/week (385 g/week). The average consumption for consumers-only is five times larger than the average consumption for all students because the denominator for consumers-only is one-fifth (20%) the size of the denominator for all students.

In our analysis of CSFII, males consumed a mean of 590 g/day of RCSD, of which 28 g—approximately 1 oz—was from a vending machine source. Females consumed a mean of 359 g/day of RCSD, of which 16 g was from a vending machine source (Table I).

In our analysis of NHANES, males consumed a mean of 668 g/day of RCSD, a 78 g increase from CSFII (Table II). Only 60 g (~2 oz) of that total was consumed at school, while 349 g was consumed at home. Females consumed a mean of 442 g/day of RCSD, an increase of 83 g from CSFII. Only 29 g of that total was consumed at school while 237 g was consumed at home. The increase in total RCSD consumption may represent a continuation of the secular trend of increased RCSD consumption that has been noted in previous studies and warrants further analysis. The differences in RCSD consumption from vending machines (CSFII) and consumed at school (NHANES) are more difficult to compare because the definitions are not comparable. For both males and females, the mean for RCSD consumed at home was more than five times that consumed at school, and the mean for RCSD consumed at home was more than half the mean total consumption. For males, the average consumption of regular fruit drinks fell by 29 g/day, from 148 g/day in CSFII to 119 g/day in NHANES. The mean consumption of regular fruit drinks was only 3 g/day lower in NHANES for females.

<table>
<thead>
<tr>
<th>Beverage</th>
<th>Total Male</th>
<th>Female</th>
<th>Vending Machine Male</th>
<th>Female</th>
<th>School Vending Machine a Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular fruit drinks/ades (g)</td>
<td>148</td>
<td>112</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Diet fruit drinks/ades (g)</td>
<td>39</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Regular carbonated soft drinks (g)</td>
<td>590</td>
<td>359</td>
<td>28</td>
<td>16</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>Diet carbonated soft drinks (g)</td>
<td>30</td>
<td>46</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Sample size</td>
<td>536</td>
<td>549</td>
<td>536</td>
<td>549</td>
<td>536</td>
<td>549</td>
</tr>
</tbody>
</table>

*aBeverages in the “School Vending Machine” category include all beverages purchased from a vending machine between the hours of 7 a.m. and 4 p.m.

Note: 370 g is equal to one 12-oz serving.

To help visualize the full distribution of RCSD consumption, we calculated cumulative density plots (Figs. 1A, 1B and 2A, 2B). Whereas a histogram shows the percentage of the sample in each arbitrarily defined category of a variable, a cumulative density plot shows the proportion of the sample that was less than or equal to every value between the minimum and the maximum. The cumulative density plot is based on the empirical cumulative density function (ECDF), which represents the proportion of the sample that is less than or equal to any given value, \( x \), of the variable. The ECDF ranges from 0 to 1. When the slope is steep, there are many respondents in that range of values, and when the slope is relatively flat, there are few respondents.

The cumulative density plots showed the consumption of RCSD on the horizontal axis and the cumulative proportion of respondents on the vertical axis. The lines for current RCSD consumption and RCSD consumption excluding school vending machines were very close. This demonstrated that removing the RCSD that may be purchased from school vending machines would make practically no difference in RCSD consumption. The results were very similar for the NHANES data (Figs. 2A and B).

### 3.2. No Association between RCSD Consumption and BMI

There was no association between RCSD consumption and BMI in the CSFII or NHANES data. Three visual techniques—the sunflower density plot, the 95% confidence ellipse, and the linear regression with associated 95% confidence interval—were used to evaluate the relationship (Figs. 3A–3D). Sunflower density plots are designed for situations in which a

---

**Table II. Mean Beverage Consumption (g/day) by Source for Males and Females 13–18 Years**

<table>
<thead>
<tr>
<th>Beverage</th>
<th>Total Male</th>
<th>Female</th>
<th>Consumed at School Male</th>
<th>Female</th>
<th>Consumed at Home Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular fruit drinks/ades (g)</td>
<td>119</td>
<td>109</td>
<td>15</td>
<td>20</td>
<td>66</td>
<td>60</td>
</tr>
<tr>
<td>Diet fruit drinks/ades (g)</td>
<td>37</td>
<td>21</td>
<td>10</td>
<td>8</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>Regular carbonated soft drinks (g)</td>
<td>668</td>
<td>442</td>
<td>60</td>
<td>29</td>
<td>349</td>
<td>237</td>
</tr>
<tr>
<td>Diet carbonated soft drinks (g)</td>
<td>38</td>
<td>54</td>
<td>2</td>
<td>1</td>
<td>23</td>
<td>33</td>
</tr>
</tbody>
</table>

Note: 370 g is equal to one 12-oz serving.

Fig. 2. (A and B) Cumulative density line plots of the original and hypothetical distributions of RCSD consumption (g/day) for males and females aged 13–18 years in NHANES removing consumption at school. The solid line represents the current distribution of RCSD consumption, and the dashed line represents the hypothetical distribution of RCSD consumption removing consumption at school. The height of each line represents the proportion of the sample that consumes less than or equal to the amount of RCSD at that point on the horizontal axis.

A traditional scatter plot cannot give a visually accurate impression of the data because many observations overlap. A sunflower density plot represents multiple overlapping points as “petals” on a “flower.” Each petal represents a single observation. In very dense areas of the plot—indicated by a darker background—each petal represents 10 observations. The sunflower density plot showed no positive association and a wide range in BMI throughout the range of RCSD consumption. The range of BMI was particularly wide at the lower levels of RCSD consumption, and nonconsumers were among the highest BMI values in the chart.

A 95% confidence ellipse contains 95% of the observations in the sample. A flat ellipse indicates no relationship, whereas a “tilted” ellipse indicates a positive or a negative association depending on the direction of the tilt. An ellipse that is vertically narrow or “skinny” indicates a smaller confidence interval for the variable on the vertical axis, and a vertically wide or “fat” ellipse indicates a large confidence interval. The confidence ellipses for RCSD and BMI were flat, indicating no relationship, and “fat,” indicating substantial variability in BMI across the range of RCSD consumption.

An unweighted linear regression found no association between RCSD consumption and BMI in NHANES. The coefficients were 0.0003 ($p > 0.39$) and −0.0001 ($p > 0.76$) for males and females, respectively. The adjusted $R^2$ for both models was 0.00. The confidence interval of the prediction was very wide, ranging from about 12.1–38.2 kg/m² for both males and females throughout the range of RCSD consumption. This implied that knowing the RCSD consumption of an individual provided no information about BMI. For example, the 95% confidence interval for the forecast of BMI for an individual male was 12.0–35.7 kg/m² for a nonconsumer and 12.6–36.4 kg/m² for a male consuming 2,220 g/day (six 12-oz servings/day). Applying sample weights and using a Taylor linearization approach to correct for sample design effects produced nearly identical results. The coefficients in the weighted and adjusted model are 0.0001 ($p > 0.74$) and 0.0002 ($p > 0.68$) for males and females, respectively. The adjusted $R^2$ of the weighted and adjusted models remains 0.00. Similar charts for CSFII also showed no association between RCSD consumption and BMI. (Figs. 3A and 3B) These results were consistent with other more sophisticated analyses of the relationship between RCSD consumption and BMI among adolescents in CSFII and NHANES 1988–1994.

3.3. Modeled BMI Distribution

The risk analysis model showed practically no difference in BMI between the original distribution of RCSD consumption and the hypothetical distribution representing the elimination of RCSD consumption from school vending machines (Table III). For males in CSFII, BMI was identical at each 10th percentile for the original and the excluded vending machine distributions of RCSD consumption. For females in CSFII, the only difference was at the 15th percentile of BMI with a value of 18.5 for the original distribution and
School Vending Machines and Risk of Adolescent Overweight

Fig. 3. (A–D) Density distribution sunflower plots with 95% confidence ellipse, linear fit regression line, and 95% confidence interval of BMI predictions for RCSD consumption (g/day) and BMI (kg/m²) for males and females aged 13–18 years using the 24-hour dietary recall from CSFII and NHANES. A single observation is represented by a single point. Multiple observations at the same point on the graph are represented as either a light or dark “flower” with “petals.” Light flowers have a clear background, and dark flowers have a gray background. Three to nine observations at the same location are represented by a clear flower with three to nine petals (● indicates three observations, + indicates four observations, □ indicates five observations, etc.). Dark flowers indicate especially dense areas of observations. Each dark flower indicates at least 10 observations at that point (dark flower with □ indicates 10 observations, dark flower with □ indicates 50 observations, etc.). Dark patches with many petals indicate areas of the plot where observations are extremely prevalent or densely clustered. The ellipse represents the area of the graph that contains 95% of the observations. The long-dashed line represents the linear regression (unweighted) of RCSD consumption on BMI, and the pair of short-dashed lines above and below the regression line represents the 95% confidence interval of the BMI prediction.

18.4 for the excluded vending machine distribution of RCSD consumption. In NHANES, the original and hypothetical BMIs were within two-tenths of a unit at each percentile for males and females (Table IV). None of the differences was statistically significant.

The cumulative density distribution graphs (Figs. 4A, 4B and 5A, 5B) illustrated how close the distributions of BMI were to one another. The distributions for current and hypothetical BMI excluding RCSD consumed at school were so close that the two lines could not be distinguished from one another—the two distributions appear to be a single curve on the graphs. The cumulative density distribution graphs were very similar for both CSFII and NHANES.

Finally, the percentage of adolescents who are overweight and underweight, as defined by the CDC age gender-specific cutoff points from the BMI-for-age growth chart, was calculated for the original and the hypothetical BMI distributions (Table V). There was practically no difference between the percent of overweight and underweight adolescents using the original and the hypothetical BMI excluding school vending machines. In the CSFII model, the percentages for females were identical, while for males
Table III. Distributions of Original BMI Versus Hypothetical BMI (kg/m²) After Excluding Vending Machine RCSD Consumption

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Males Original</th>
<th>Males No Vending</th>
<th>Females Original</th>
<th>Females No Vending</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>17.0</td>
<td>17.0</td>
<td>16.7</td>
<td>16.7</td>
</tr>
<tr>
<td>15</td>
<td>18.6</td>
<td>18.6</td>
<td>18.5</td>
<td>18.4</td>
</tr>
<tr>
<td>25</td>
<td>19.7</td>
<td>19.7</td>
<td>19.4</td>
<td>19.4</td>
</tr>
<tr>
<td>35</td>
<td>20.4</td>
<td>20.4</td>
<td>20.1</td>
<td>20.1</td>
</tr>
<tr>
<td>45</td>
<td>21.2</td>
<td>21.2</td>
<td>20.8</td>
<td>20.8</td>
</tr>
<tr>
<td>55</td>
<td>22.2</td>
<td>22.2</td>
<td>21.5</td>
<td>21.5</td>
</tr>
<tr>
<td>65</td>
<td>23.1</td>
<td>23.1</td>
<td>22.5</td>
<td>22.5</td>
</tr>
<tr>
<td>75</td>
<td>24.7</td>
<td>24.7</td>
<td>24.0</td>
<td>24.0</td>
</tr>
<tr>
<td>85</td>
<td>26.6</td>
<td>26.6</td>
<td>26.4</td>
<td>26.4</td>
</tr>
<tr>
<td>95</td>
<td>30.3</td>
<td>30.3</td>
<td>30.9</td>
<td>30.9</td>
</tr>
</tbody>
</table>


The results were similar for the NHANES model. For males, the percentage overweight decreased by 0.4 because two males moved from the overweight to the recommended weight category in the hypothetical. The percent of males who were underweight was unchanged when school consumption was excluded. For females, the percentages overweight and underweight were identical in the original and the hypothetical.

The analysis of the simulated data that randomly sampled consumption for 365 days also showed little difference in the BMI distributions. At each 10th percentile, the hypothetical BMI excluding school consumption was either equal to or one-tenth kg/m² lower than the original BMI (Table VI). The cumulative density plots showed practically no difference in BMI distributions between the original distribution and the hypothetical BMI distributions (not shown).

Table IV. Distributions of Original BMI Versus Hypothetical BMI (kg/m²) After Excluding RCSD Consumption at School

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Males Original</th>
<th>Males No School</th>
<th>Females Original</th>
<th>Females No School</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>17.0</td>
<td>17.0</td>
<td>17.8</td>
<td>17.7</td>
</tr>
<tr>
<td>15</td>
<td>18.7</td>
<td>18.7</td>
<td>19.1</td>
<td>19.1</td>
</tr>
<tr>
<td>25</td>
<td>20.0</td>
<td>19.9</td>
<td>20.1</td>
<td>20.1</td>
</tr>
<tr>
<td>35</td>
<td>20.8</td>
<td>20.8</td>
<td>21.1</td>
<td>21.1</td>
</tr>
<tr>
<td>45</td>
<td>21.8</td>
<td>21.7</td>
<td>22.3</td>
<td>22.2</td>
</tr>
<tr>
<td>55</td>
<td>23.1</td>
<td>23.1</td>
<td>23.4</td>
<td>23.4</td>
</tr>
<tr>
<td>65</td>
<td>24.5</td>
<td>24.5</td>
<td>24.9</td>
<td>24.9</td>
</tr>
<tr>
<td>75</td>
<td>26.9</td>
<td>26.9</td>
<td>27.0</td>
<td>27.0</td>
</tr>
<tr>
<td>85</td>
<td>29.8</td>
<td>29.8</td>
<td>30.4</td>
<td>30.4</td>
</tr>
<tr>
<td>95</td>
<td>36.1</td>
<td>36.1</td>
<td>35.8</td>
<td>35.7</td>
</tr>
</tbody>
</table>


Fig. 4. (A and B) Cumulative density line plots of the original and hypothetical distributions of BMI (kg/m²) for males and females aged 13–18 years in CSFII removing RCSD consumption from school vending machines. The solid line represents the original distribution of BMI, and the dashed line represents the hypothetical distribution of BMI after subtracting 0.24 times RCSD consumption (370 g/serving) from school vending machines. The height of each line represents the proportion of the sample that is less than or equal to the BMI value at that point on the horizontal axis. The distributions are so close that the dashed line representing the hypothetical BMI cannot be distinguished from the solid line representing the original BMI.

3.4. Sensitivity Tests

Sensitivity tests were performed to examine the impact of the dose-response estimation and the exposure estimation on the distribution of BMI using the NHANES data. Sensitivity Test 1 used a dose-response estimation from Berkey et al., which
School Vending Machines and Risk of Adolescent Overweight

Fig. 5. (A and B) Cumulative density line plots of the original and hypothetical distributions of BMI (kg/m²) for males and females aged 13–18 years in NHANES removing RCSD consumption at school. The solid line represents the original distribution of BMI, and the dashed line represents the hypothetical distribution of BMI after subtracting 0.24 times RCSD consumption (370 g/serving) at school. The height of each line represents the proportion of the sample that is less than or equal to the BMI value at that point on the horizontal axis. The distributions are so close that the dashed line representing the hypothetical BMI cannot be distinguished from the solid line representing the original BMI.

Table V. Proportion Overweight and Underweight After Hypothetical Reductions of RCSD Consumption

<table>
<thead>
<tr>
<th>RCSD Condition</th>
<th>Males 13–18 Years</th>
<th>Females 13–18 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% Overweight</td>
<td>% Underweight</td>
</tr>
<tr>
<td>Original</td>
<td>27.3</td>
<td>5.5</td>
</tr>
<tr>
<td>No school vending</td>
<td>27.1</td>
<td>5.5</td>
</tr>
<tr>
<td>Original</td>
<td>28.5</td>
<td>5.9</td>
</tr>
<tr>
<td>No school consumption</td>
<td>28.1</td>
<td>5.9</td>
</tr>
</tbody>
</table>

Analysis by authors.

estimated an association of 0.02 kg/m² per one serving change in sugar-sweetened beverage consumption. This roughly one order of magnitude lower than the dose response of 0.24 used in our base model. Sensitivity Test 2 used an exposure estimation based on an 8-oz serving size rather than a 12-oz serving size. This effectively increased exposure by one-third for consumers. Sensitivity Test 3 also used an 8-oz serving size and increased the dose-response estimation to 1.0 kg/m² per one serving change—a dose response that is more than four times as large as any reported in the literature.

Sensitivity Test 1 showed even less impact on the BMI distribution than did the base model. The percentiles for the original BMI and the hypothetical BMI removing consumption from schools are identical with the exception that the 45th percentile for females is 0.1 kg/m² lower after RCSD consumption at school has been removed compared to the original BMI. Sensitivity Test 2 used a higher exposure estimate and showed a slightly larger impact on BMI for males than did our base model. Three of the percentiles for hypothetical BMI were 0.1 kg/m² lower in Sensitivity Test 2 than in the base model. The difference between Sensitivity Test 2 and the base model was less than 0.1 kg/m² for each percentile for females. Sensitivity Test 3—which used a dose response four times larger than any identified in the literature and the higher exposure estimate—showed an impact on the BMI distributions. The hypothetical BMI distribution was lower in most of the percentiles, usually 0.1–0.4 units. The cumulative density graphs diverged slightly, indicating a small shift in the BMI distribution.

4. DISCUSSION

Some suggest that removing RCSD from vending machines at school will decrease consumption and
<table>
<thead>
<tr>
<th>Percentile</th>
<th>Males Original</th>
<th>Males No School</th>
<th>Females Original</th>
<th>Females No School</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>16.8</td>
<td>16.8</td>
<td>17.9</td>
<td>17.8</td>
</tr>
<tr>
<td>15</td>
<td>18.7</td>
<td>18.6</td>
<td>19.1</td>
<td>19.1</td>
</tr>
<tr>
<td>25</td>
<td>20.0</td>
<td>20.0</td>
<td>21.1</td>
<td>21.1</td>
</tr>
<tr>
<td>35</td>
<td>20.8</td>
<td>20.7</td>
<td>22.3</td>
<td>22.2</td>
</tr>
<tr>
<td>45</td>
<td>21.8</td>
<td>21.7</td>
<td>23.4</td>
<td>23.4</td>
</tr>
<tr>
<td>55</td>
<td>23.0</td>
<td>23.0</td>
<td>24.9</td>
<td>24.8</td>
</tr>
<tr>
<td>65</td>
<td>24.5</td>
<td>24.4</td>
<td>26.8</td>
<td>26.8</td>
</tr>
<tr>
<td>75</td>
<td>26.9</td>
<td>26.9</td>
<td>30.1</td>
<td>30.1</td>
</tr>
<tr>
<td>85</td>
<td>30.1</td>
<td>30.0</td>
<td>35.8</td>
<td>35.7</td>
</tr>
</tbody>
</table>

Source: Data: Males 13–18 years, Monte Carlo bootstrap simulation based on NHANES. Analysis by authors.

subsequently help address overweight among adolescents. The findings from this risk assessment indicate that removing vending machines from schools or eliminating RCSD from the vending machines will have little effect on RCSD consumption and practically no effect on BMI or overweight. Even the upper plausible estimate of the impact of RCSD consumption from school vending machines on BMI is minimal. Eliminating RCSD consumption at school only shifted the BMI distribution by $1/10$ kg/m² at a few percentiles.

It is important to emphasize that all of the estimates presented in the base model in this article reflect the upper end of the possible impact of RCSD on BMI and overweight. We chose the largest estimates of consumption of RCSD from vending machines from nationally representative data sets. We chose the largest estimate of the relationship between RCSD consumption and BMI that we identified in the literature for our base model. In addition, we assumed that there would be no replacement of RCSD from school vending machines from other sources. The actual impact of removing school vending machines or RCSD from school vending machines on BMI would almost certainly be less than the estimates presented.

There are two reasons why there is so little difference between the current distribution of BMI and the modeled distribution of BMI eliminating RCSD from school vending machines. First, RCSD consumption from school vending machines is a small fraction of total RCSD consumption. Estimates of RCSD consumption from school vending machines from the three surveys used in our analysis ranged from a low of 15 g/day in NFO-SIP 2001–2002 to a high of 60 g/day in NHANES 1999–2000. The NHANES estimate was likely an overestimate of RCSD consumption from vending machines since it included all RCSD consumed at school, whether from vending machines or other sources.

Second, the relationship between RCSD consumption and BMI does not appear to be strong. Our analysis of the relationship between RCSD consumption and BMI in CSFII and NHANES showed no relationship. The estimated coefficients were nearly zero, the $p$-values were large, and the adjusted $R^2$ was 0.00. The findings in the scientific literature have been mixed. A 19-month prospective observational study with 548 school children aged 11–12 years living in Massachusetts reported an average increase in BMI of 0.24 kg/m² over 19 months for each serving per day increase in sugar-sweetened beverage consumption ($p = 0.03$). This implies that consumption of sugar-sweetened beverages must be reduced by more than 4 servings/day to reduce BMI by 1 kg/m². Given that average RCSD consumption is less than 2 servings/day for males and about 1.2 servings/day for females, the magnitude of the relationship between RCSD consumption and BMI does not provide enough leverage to make a meaningful difference in BMI distribution of the population. Furthermore, the average increase in sugar-sweetened beverage consumption over the course of the study was 0.22 servings/day, from 1.22 at baseline to 1.44 at follow-up 19 months later. A 0.22 servings/day increase would predict a 0.05 kg/m² increase in BMI [0.22 servings/day $\times 0.24$ kg/m² per serving/day].

A prospective observational study of more than 10,000 boys and girls aged 9–14 years (U.S. Growing Up Today Study or GUTS) found an average increase of 0.03 kg/m² per serving/day over a 1-year time period for males ($p = 0.04$). For females, there was a nonsignificant increase of 0.02 kg/m² per serving/day ($p = 0.096$). The coefficients are much smaller and not statistically significant once total energy was introduced as a control variable.

A separate analysis of GUTS found no association between consumption of snack foods and annual change in BMI z-score. The estimated coefficients were negative, small, and not significant for both boys and girls. Adding sugar-sweetened beverages to the snack food category “did not meaningfully change the results.”

A cluster, randomized controlled trial that was designed to discourage consumption of “fizzy” drinks...
among 7- to 11-year-old British school children showed that consumption of carbonated drinks with sugar was unchanged in the control group and decreased by 0.3 of a (250 ml) serving over a 3-day period (25 ml/day or <1 oz/day) in the intervention group. Neither change was statistically significant nor was there a statistically significant difference between the control and intervention groups. Furthermore, the mean change in BMI was 0.8 kg/m² for the control group and 0.7 kg/m² for the intervention group. This difference was not statistically significant.\\(^{(11)}\)

A prospective cohort study of 1,345 children aged 2–5 years who were participating in the North Dakota Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) found no statistically significant association between beverage consumption and change in either weight or BMI. Specifically with regard to RCSD consumption, the estimated coefficient in the multivariate adjusted model for weight was $−0.00 ± 0.04$ ($p = 0.95$) and for BMI was $−0.01 ± 0.02$ ($p = 0.58$).\\(^{(12)}\)

This brief review of the literature does not constitute a comprehensive meta-analysis and does not prove that RCSD consumption is not associated with BMI. Each study has limitations, particularly the measurement error inherent in self-reported dietary data. The point is that many well-designed studies using validated dietary recall instruments have found an association between RCSD consumption and BMI that is close to zero. We chose the largest of these estimates for the dose-response estimation in the base model.

The elimination of school vending machines would have certain costs. The most obvious cost is the lost revenue to the schools from sales of products from vending machines. In a recent focus group study, parents emphasized the importance of this revenue and commented that they would rather the money from soft drink sales go to schools than to stores or gas stations. In the same study, parents also emphasized the importance of control and choice for students. The main theme from the focus groups was that high school students need to start making their own decisions and taking responsibility for their beverage choices.\\(^{(40)}\)

A much larger survey of attitudes toward obesity policy also found that Americans express relatively little support for obesity-targeted policies.\\(^{(41)}\) More recently, a survey conducted in 2004 showed that 41.3% of adults viewed childhood obesity as a “very serious” problem. Childhood obesity was cited by a smaller percentage of the respondents than any of the other five problems included on the survey, specifically underage smoking (42.4%), teen sex (44.3%), underage drinking (47.4%), youth violence (50.4%), and drug abuse (54.5%). Parents were almost universally viewed as having “a lot of responsibility” to reduce childhood obesity (90.7%). Far fewer respondents said schools (30.0%) and government (16.8%) had “a lot of responsibility.” Among specific interventions, 93.9% favored “requiring schools to teach students healthy eating and exercise habits” and 82.3% favored “requiring more physical education classes in school.” Restricting sales from school vending machines was favored by 85.4% of respondents, but removing all vending machines was supported by only 35.9%. In general, the authors found public support for strategies that offer health information, limit unhealthy food promotion, and increase healthy nutrition and physical activity choices, but the public was generally opposed to regulatory and tax- or cost-based interventions.\\(^{(42)}\)

Our results by no means preclude possible benefits to some individuals if RCSD sales in secondary schools are reduced or banned. Further epidemiological studies and controlled trials of intervention programs are needed to develop appropriate and effective policies and interventions to reduce the risk of overweight among adolescents. What we can say is that even upper-end estimates, which would correspond to upper percentiles of the reduction in BMI resulting from elimination of school-related intakes of RCSD are very modest, and, in our view, do not support a policy that bans sales of RCSD in secondary schools. Such a policy would have a very modest impact on childhood obesity and measurable costs in terms of revenue reduction and less personal choice for students and adults using school facilities.

Finally, we note that our analysis was limited to the direct impact of RCSD purchased or consumed in schools on BMI and risk of overweight. We did not evaluate other health endpoints, and we did not evaluate any social or indirect impacts associated with making branded beverages available in schools.

5. CONCLUSION

The National Research Council has pointed out that in a democracy the government and experts should take great care in how they manage risk. Ideally, risks should be managed by providing appropriate information and resources without undue advocacy so that citizens can make appropriate decisions to deal with risk.\\(^{(43)}\) Our findings suggest that focusing
on RCSD from vending machines in schools will not have a significant impact on BMI.

Our analysis also demonstrates that the principles of risk analysis can be successfully applied to the field of nutrition and public health, and we hope that this example will encourage further applications of risk analysis to important questions in nutrition and public health. Risk analysis has proven to be an important tool in diverse fields, including automotive safety, environmental pollution, drug safety, pesticides in the diets of children, and food safety. The principles of risk analysis can help to integrate data from many sources, prioritize research needs by identifying the most critical research gaps, and provide important input to policy and regulatory decisions about which proposed interventions are most likely to yield the largest reduction in risk. Nutrition and public health can gain the same benefits by adopting these principles.

ACKNOWLEDGMENT

This project was supported by an unrestricted gift from the American Beverage Association.

REFERENCES


