Association Between Intake of Added Sugars with Nutrient Intakes for Children and Adolescents Ages 9-18 years old

Brenna Curley

Iowa State University

Joint Statistical Meetings
August 6, 2013
Outline

▶ Question of Interest

▶ Previous Work

▶ The Data
  ▶ Definition of Empty Calories
  ▶ Summary Statistics
  ▶ Model Setup

▶ Methodology
  ▶ Measurement Error Model

▶ Results and Discussion
Question of Interest

Goals:

- To explore any associations that may exist between consumption of added sugars and nutrient intakes for children and adolescents ages 9-18y.

- Longterm average intakes are of interest, but only 2 days data collected in surveys so want to minimize nuisance day-to-day variation.
Previous Work

Conflicting results – no clear negative association found yet. Previous debates on:

- Use of ratios
- Accounting for total energy

Our model:

- Use caloric densities
- Account for other empty calories and sources of energy
- Scale responses by total energy
- Interest is in usual intakes so use a measurement error model
The Data

- National Health and Nutrition Examination Survey (NHANES) 2003-2008
- US Department of Agriculture food conversion table to calculate nutrient intakes given food items
- MyPyramid Equivalents Database — information on intakes of added sugar and discretionary fat
- 2 independent days of data per individual
The Data — Definitions

- **Added Sugars**
  - Eaten separately, an ingredient in processed or prepared food, or added to foods after they are prepared
  - *Examples*: Sugars in cakes, chocolates, or jams
  - Sugar substitutes, fruit juice concentrates in recipes, fructose in fruit
  - NOT considered added sugars

- **Discretionary (Solid) Fats**
  - Fat in excess of the allowable amount
  - *Example*: Any fat in whole milk that is not in skim milk is discretionary
  - Solid fats are fats that are solids at room temperature
### The Data — Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>9-13y Male</th>
<th>14-18y Male</th>
<th>9-13y Female</th>
<th>14-18y Female</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BMI</strong></td>
<td>20.08</td>
<td>23.36</td>
<td>20.31</td>
<td>23.50</td>
</tr>
<tr>
<td><strong>Total kcal</strong></td>
<td>2174.24</td>
<td>2584.37</td>
<td>1922.87</td>
<td>1872.35</td>
</tr>
<tr>
<td><strong>Added Sugar</strong></td>
<td>417.80</td>
<td>391.41</td>
<td>402.54</td>
<td>357.03</td>
</tr>
<tr>
<td>(%E&lt;sub&gt;AS&lt;/sub&gt;)</td>
<td>(19.87)</td>
<td>(15.79)</td>
<td>(21.65)</td>
<td>(19.61)</td>
</tr>
<tr>
<td><strong>Solid Fats</strong></td>
<td>597.11</td>
<td>579.07</td>
<td>590.26</td>
<td>542.75</td>
</tr>
<tr>
<td>(%E&lt;sub&gt;SOL&lt;/sub&gt;)</td>
<td>(28.21)</td>
<td>(23.07)</td>
<td>(31.47)</td>
<td>(30.47)</td>
</tr>
</tbody>
</table>

**Table:** Overall means $\bar{X}_{\pi_i}$ and $\bar{Y}_{\pi_i}$ for all variables of interest taking into account survey weights, $\pi_i$, for each individual $i$. Added sugar and solid fat intakes measured in their original units of kcal (and as a ratio with total energy, multiplied by 100).
The Data — Model Setup

Variables of interest:

Responses:
- Calcium (mg), Fiber (g), Vitamin E (mg), Magnesium (mg), Phosphorus (mg), and Potassium (mg) as nutrient densities.
- Natural log transformations were taken for more symmetrical distributions.

Predictors:
- Added Sugar (%$E_{AS}$) and Discretionary Solid Fat (%$E_{SOL}$)
- Age (years), BMI, weekend and race/ethnicity indicators
  - Consider separate age-sex groups: Males and Females 9-13y, Males and Females 14-18y
Setup:

- Only 2 days data per individual
- Interest is in the longterm average intakes
- Will use a measurement error model (Fuller 1987) to reduce bias in estimated regression coefficients attributed to nuisance day-to-day variation
Methodology - Measurement Error Model

Going to let $Y_{ij}$ be the observed nutrient intake for the $i$th individual on the $j$th day.

But we are interested in $y_i$, the **unobservable usual** (long term average) intake of a nutrient.

Let:

$$Y_{ij} = y_i + w_{ij}$$ 

$$y_i \sim N(\mu_y, \sigma_{yy})$$ 

$$w_{ij} \sim N(0, \sigma_{ww})$$

where $i=1,...,n$ individuals and $j=1,2$ days per individual. The values $\sigma_{yy}$ and $\sigma_{ww}$ denote the between and within person variances and $y_i$ and $w_{ij}$ assumed independent.
Methodology — Measurement Error Model

Both the response and predictors have measurement error attached to the observed intakes.

To simplify we will look at the case with just one predictor, giving us the model:

$$y_i = \beta_0 + x_i\beta_1 + q_i,$$

for

$$Y_{ij} = y_i + w_{ij}$$
$$X_{ij} = x_i + u_{ij},$$

where $E(Y_{ij}|i) = y_i$ and $E(X_{ij}|i) = x_i$. 
Hence we get the model:

\[ Y_{ij} = \beta_0 + x_i \beta_1 + e_{ij}, \]

where \( e_{ij} = q_i + w_{ij} \) is composed of two independent parts:

- \( q_i \): The error in measuring \( y_i \) since observations vary from day to day
- \( w_{ij} \): The error in the equation when measuring usual intake

We also note the possible correlation between our observed response and predictor variables such that:

\[
\begin{bmatrix}
    w_{ij} \\
    u_{ij}
\end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix}
    \sigma_{ww} & \sigma_{wu} \\
    \sigma_{wu} & \sigma_{uu}
\end{bmatrix} \right).
\]
Methodology — Linear Regression Estimates

Regression coefficients are adjusted to account for:

- The error attached to our observed, measured daily intakes for both response and predictor variables
- Any correlation between response and predictor variables
- Survey weights ($\pi_i$)
Thus we calculate our regression coefficients as:

\[
\tilde{\beta}_1 = \frac{M_{X \pi Y} - \hat{\sigma}_{wu}}{M_{X \pi X} - \hat{\sigma}_{uu}}; \quad \tilde{\beta}_0 = \bar{Y}_{\pi..} - \bar{X}_{\pi..}\tilde{\beta}_1
\]

Where:

- \(M_{X \pi X}\) and \(M_{X \pi Y}\) be the variance and covariance, respectively, of the averaged, usual intakes per individual

- \(\hat{\sigma}_{wu}\) and \(\hat{\sigma}_{uu}\) are the weighted averaged variance and covariances of intakes for individuals — A robust estimate is computed using weighted medians to avoid poor model fit.

- \(\bar{Y}_{\pi..}\) and \(\bar{X}_{\pi..}\) are the weighted overall averages of the response and predictor variables.
In order to proceed with any inference about our coefficients we must consider the variance of our estimates.

Recall we must account for both:

- Measurement Error
- Survey Weights
Methodology — Variance of Coefficients

Thus, we use a fractional jackknife estimate:

$$\hat{V}_{FJK}(\tilde{\beta}_1) = \sum_{h=1}^{H} \frac{n_h - 1}{n_h} \sum_{j=i}^{n_h} (\tilde{\beta}_{(hj)} - \tilde{\beta}_{full})(\tilde{\beta}_{(hj)} - \tilde{\beta}_{full})$$

where:

- $n_h$ is the number of PSU’s in stratum $h$ — two in our case.

- The estimate $\tilde{\beta}_{full}$ is computed with the original weights

- The estimates $\tilde{\beta}_{(hj)}$ are computed using adjusted weights such that individuals in PSU $j$ of stratum $h$ are weighted lower than those in stratum $h$ but not PSU $j$; weights not in stratum $h$ are left alone.
Results — Correlations

**Added Sugar Correlations**

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>Male 9-13y</th>
<th>Male 14-18y</th>
<th>Female 9-13y</th>
<th>Female 14-18y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcium</td>
<td>−0.15</td>
<td>−0.21</td>
<td>−0.15</td>
<td>−0.13</td>
</tr>
<tr>
<td>Fiber</td>
<td>−0.13</td>
<td>−0.09</td>
<td>−0.11</td>
<td>−0.07</td>
</tr>
<tr>
<td>Vitamin E</td>
<td>−0.24</td>
<td>−0.10</td>
<td>−0.16</td>
<td>−0.12</td>
</tr>
<tr>
<td>Magnesium</td>
<td>−0.29</td>
<td>−0.19</td>
<td>−0.11</td>
<td>−0.07</td>
</tr>
<tr>
<td>Phosphorus</td>
<td>−0.43</td>
<td>−0.30</td>
<td>−0.17</td>
<td>−0.28</td>
</tr>
<tr>
<td>Potassium</td>
<td>−0.32</td>
<td>−0.17</td>
<td>−0.25</td>
<td>−0.14</td>
</tr>
</tbody>
</table>

**Table:** Added sugar correlations after accounting for survey weights as well as measurement error. Correlations with corresponding significant parameter estimates (at $\alpha = .05$ level) are shaded.
We used a model that accounted for:

- Total Energy
- Measurement Error
- Survey Weights

Found results supporting a negative association between added sugar and intakes of certain essential nutrients for individuals ages 9-18y in the U.S.
References

Acknowledgments:

Special thanks to Alicia Carriquiry (ISU).

Also thanks to Ulrike Genschel (ISU), Helen Jensen (ISU), and Patricia Guenther (CDC) for their input.

Contact Info: curleyb@iastate.edu