

Dietary Screener in CHIS 2009: Variance Adjustment

Introduction

Dietary intake estimates from the California Health Interview Survey (CHIS) Dietary Screener are rough estimates of usual intake of fruits and vegetables and added sugars. They are less accurate than more detailed methods (e.g. 24-hour recalls). However, validation research suggests that the estimates may be useful to characterize a population's median intakes, to discriminate among individuals or populations with regard to higher vs. lower intakes, to track dietary changes in individuals or populations over time, and to allow examination of interrelationships between diet and other variables. In addition, dietary estimates from the CHIS could be used to augment national data using similar methods.

What is the variance adjustment estimate and why is it needed?

Data from the CHIS Dietary Screener are individuals' reports about their intake and, like all self-reports, contain some error. The algorithms we use to estimate servings of fruits and vegetables and added sugars calibrate the data to 24-hour recalls. The screener estimate of intake represents what we expect the person would have reported on his or her 24-hour recall, given what s/he reported on the individual items in the screener. As a result, the mean of the screener estimate of intake should equal the mean of the 24-hour recall estimate of intake in the population. (It would also equal the mean of true intake in the population if the 24-hour recalls were unbiased. However, many studies suggest that recalls underestimate individuals' true intakes).

When describing a population's distribution of dietary intakes, the parameters needed are an estimate of central tendency (i.e., mean or median) and an estimate of spread (i.e., variance). The variance of the screener, however, is expected to be smaller than the variance of true intake because the screener prediction formula estimates the conditional expectation of true intake given the screener responses, and in general, the variance of a conditional expectation of a variable X is smaller than the variance of X itself.

As a result, the screener estimates of intake cannot be used to estimate quantiles (other than median) or prevalence estimates of true intake unless they are first adjusted so that they have approximately the same variance as true intake.

When is it appropriate to use variance adjustment estimates?

The appropriate use of the screener information depends on the analytical objective. Following are suggested procedures for various analytical objectives.

Analytical Objective	Procedure
Estimate mean or median intake in the population or within subpopulations.	Use the unadjusted screener estimate of intake.
Estimate quantiles (other than median) of the distribution of intake in the population; estimate	Use the variance-adjusted screener estimate.

prevalence of attaining certain levels of dietary intake.	
Classify individuals into exposure categories (e.g., meeting recommended intake vs. not meeting recommended intake) for later use in a regression model.	Use the variance-adjusted screener estimates to determine appropriate classification into categories.
Use the screener estimate as a continuous covariate in a multivariate regression model.	Use the unadjusted screener estimate.

How were the variance adjustment factors estimated?

We developed procedures to estimate the variance of true intake using data from 24-hour recalls, by taking into consideration within-person variability.^{1,2} We extended these procedures to allow estimation of the variance of true intake using data from the screener. The resulting variance adjustment factors adjust the screener variance to approximate the variance of true intake in the population.

We used two external validation datasets to estimate the adjustment factors: the Eating at America's Table Study (EATS) and the Observing Protein and Energy Nutrition Study (OPEN). The results indicate that the adjustment factors differ by gender for each dietary variable. Under the assumption that the variance adjustment factors appropriate to the California Health Interview Survey are similar to those in these external studies, the variance-adjusted screener estimates of intake should have variances closer to the estimated variance of true intake that would have been obtained from repeat 24-hour recalls.

Dietary Variable	Variance Adjustment Factors	
	Men	Women
Fruits and vegetables without dried beans (cup equivalents)	1.72	1.37
Fruits and vegetables without French fries and dried beans (cup equivalents)	1.73	1.39
Added sugars (tsp)	1.26	1.28

How are the variance adjustment factors applied?

The screener predicts intake on a transformed scale (i.e., the square root of cup equivalents of fruits and vegetables and the cube-root of teaspoons of added sugars). The variance adjustment factor is applied to predicted intake on the transformed scale. The results can then be back-transformed to obtain estimates in the original units.

Adjust the screener estimate of intake by:

- multiplying intake by an adjustment factor (an estimate of the ratio of the standard deviation of true intake to the standard deviation of screener intake); and
- adding a constant so that the overall mean is unchanged.

The formula for the variance-adjusted screener is:

$$\text{variance-adjusted screener} = (\text{variance adjustment factor}) * (\text{unadjusted screener} - \text{mean}_{\text{unadj scr}}) + \text{mean}_{\text{unadj scr}}$$

A similar variance adjustment procedure is used to estimate prevalence of intakes for the 2000 NHIS in:

Thompson FE, Midthune D, Subar AF, McNeel T, Berrigan D, Kipnis V. Dietary intake estimates in the National Health Interview Survey, 2000: Methodology, results, and interpretation. *J Am Dietet Assoc* 2005;105:352-63. [[View Abstract](#)]

The following variance-adjusted variables are available for CHIS 2009:

Variable Name	Label
FVNB2ADJ	VARIANCE-ADJ DAILY CUP EQUIV OF FRUITS/VEG EXCL BEANS
FVNFB2AJ	VARIANCE-ADJ DAILY CUP EQUIV FRUITS/VEG EXCL FRENCH FRIES & BEANS
SUG2_ADJ	VARIANCE-ADJ DAILY TEASPOONS OF ADDED SUGAR

Attenuation of Regression Parameters Using Screener Estimates

When the screener estimate of dietary intake is used as a continuous covariate in a multivariate regression, the estimated regression coefficient will typically be attenuated (biased toward zero) due to measurement error in the screener. This "attenuation factor"³ can be estimated in a calibration study and used to deattenuate the estimated regression coefficient (by dividing the estimated regression coefficient by the attenuation factor).

We estimated attenuation factors in the EATS and OPEN data (see the following table).

Dietary Variable	Attenuation factors for screener-predicted intake	
	Men	Women
(Square-root) Fruits and vegetables without dried beans (cup equivalents)	1.00	0.82
(Square-root) Fruits and vegetables without French fries and dried beans (cup equivalents)	1.02	0.87
(Cube-root) Added sugars	0.80	0.86

If the screener values are categorized into quantiles and the resulting categorical variable is used in a linear or logistic regression, the bias (due to misclassification) is more complicated because the categorization can lead to differential misclassification in the screener⁴. Although methods may be available to correct for this^{5,6}, it is not simple, nor are we comfortable suggesting how to do it at this time.

Even though the estimated regression coefficients are biased (due to measurement error in the screener or misclassification in the categorized screener), tests of whether the regression coefficient is different from zero are still valid. For example, if one used the SUDAAN REGRESS procedure with fruit and vegetable intake (estimated by the screener) as a covariate in the model, one could use the Wald F statistic provided by SUDAAN to test whether the regression coefficient was statistically significantly different from zero. This assumes that only one covariate in the model is measured with error. When multiple covariates are measured with error, the Wald F test that a single regression coefficient is zero may not be valid, although the test that the regression coefficients for all covariates measured with error are zero is still valid.

References

1. National Research Council. Nutrient Adequacy: Assessment Using Food Consumption Surveys. Washington, DC: National Academy Press, 1986. [\[More information about this publication\]](#)
2. Institute of Medicine. Dietary Reference Intakes: Applications in Dietary Assessment. Washington, DC: National Academy Press, 2000. [\[More information about this publication\]](#)
3. Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat Med* 1989;8:1051-69. [\[View Abstract\]](#)
4. Flegal KM, Keyl PM, Nieto FJ. Differential misclassification arising from nondifferential errors in exposure measurement. *Am J Epidemiol* 1991;134:1233-44. [\[View Abstract\]](#)

5. Flegal KM, Brownie C, Haas JD. The effects of exposure misclassification on estimates of relative risk. *Am J Epidemiol* 1986;123:736-51. [\[View Abstract\]](#)
6. Morrissey MJ, Spiegelman D. Matrix methods for estimating odds ratios with misclassified exposure data: extensions and comparisons. *Biometrics* 1999;55:338-44. [\[View Abstract\]](#)