

# measurement ERROR webinar series

## Combining self-report dietary intake data and biomarker data to reduce the effects of measurement error

(Webinar 11)

### Objectives:

- To describe the motivation for combining dietary self-reports and biomarkers.
- To provide an overview of different methods of combining self-reports and biomarkers, their aims, and the knowledge required for implementing each method.
- To identify the potential gains of such combinations and the limitations of the methods

### Recommended resources:

- Freedman LS, Kipnis V, Schatzkin A, Tasevska N, Potischman N. Can we use biomarkers in combination with self-reports to strengthen the analysis of nutritional epidemiologic studies? *Epidemiol Perspect Innov.* 2010;7(1):2.
- Freedman LS, Midthune D, Carroll RJ, Tasevska N, Schatzkin A, Mares J, Tinker L, Potischman N, Kipnis V. Using regression calibration equations that combine self-reported intake and biomarker measures to obtain unbiased estimates and more powerful tests of dietary associations. *Am J Epidemiol.* 2011 Nov 1. [Epub ahead of print]
- Freedman LS, Tasevska N, Kipnis V, Schatzkin A, Mares J, Tinker L, Potischman N. Gains in statistical power from using a dietary biomarker in combination with self-reported intake to strengthen the analysis of a diet-disease association: an example from CAREDS. *Am J Epidemiol.* 2010;172(7):836-42.

### Key terms:

<b>Attenuation</b>	Bias of the estimated regression coefficient in the direction of zero due to measurement error in a covariate; bias to the null.
<b>Attenuation factor</b>	The multiplicative factor by which an estimate of a regression coefficient is shrunk due to measurement error in a covariate.
<b>Biomarker</b>	For the purposes of the webinar series, a biological (usually biochemical) indicator or measure of dietary intake or nutritional status.
<b>Calibration equation</b>	An equation for predicting a true covariate value (for example, usual dietary intake) given all of the observed covariates in a regression model; usually developed from data gathered in a calibration substudy.

<b>Calibration substudy</b>	A small-scale study performed to enable calibration of the main study instrument using a reference instrument; data from the substudy are used as the basis for regression calibration. Such studies can be conducted either as external calibration or internal calibration.
<b>Carotenoids in Age-Related Eye Disease Study (CAREDS)</b>	Ancillary study of the Women’s Health Initiative (WHI) Observational Study. WHI was a prospective study of 93,676 postmenopausal women aged 50-70 years at time of enrollment (1994-1998).
<b>Case-control study</b>	A type of study that classifies individuals with regard to current disease status (as cases or controls) and relates this to past (retrospectively reported) exposures.
<b>Causal/Causation</b>	A type of relationship between two variables in which a change in the value of one causes the value of the other to change.
<b>Classical measurement error</b>	A type of measurement error consisting of random within-person error, which has a mean of zero and constant variance and which is independent of the true value.
<b>Concentration biomarker</b>	A marker of the concentration of a specific chemical or compound in blood, urine, or tissues that is subject to substantial interindividual differences in metabolism; related to and can be used as an indirect measure of dietary intake.
<b>Confounding</b>	Distortion of an association between an exposure and a health outcome by a third variable that is related to both.
<b>Explanatory variable</b>	A variable thought to be related to an outcome in a regression model.
<b>Food frequency questionnaire (FFQ)</b>	A dietary instrument that asks respondents to report their usual frequency of consumption of each food in a list of foods over a specific period of time.
<b>Logistic regression</b>	Statistical model that relates a binary outcome to one or more independent variables, using the logit link.
<b>Measurement error</b>	The difference between the observed or measured value and the true value.

<b>Mediation</b>	A phenomenon by which the causal effect of an exposure on an outcome is partially or wholly obtained through its influence on a third variable (the intermediate variable), which in turn affects the outcome. The intermediate variable is said to mediate the effect of the exposure on the outcome.
<b>Odds ratio</b>	A statistical measure that quantifies the association between an exposure and a health outcome; often used in case-control studies.
<b>Outcome</b>	The target variable; also referred to as the dependent variable in a regression model; often a health outcome, such as the occurrence of a specified disease.
<b>Power</b>	The probability that a test correctly rejects the null hypothesis when the alternative hypothesis is true.
<b>Recovery biomarker</b>	Specific biologic products that are directly related to intake and not subject to homeostasis or substantial interindividual differences in metabolism; for example, doubly labeled water for energy intake and urinary nitrogen for protein intake.
<b>Regression calibration</b>	A statistical method for correcting estimated regression coefficients for bias due to measurement error in one or more continuous covariates.
<b>Regression model</b>	A model used to quantify a relationship between an outcome and one or more explanatory variables; such models are used to estimate usual intake and relate it to other variables of interest.
<b>Standard deviation</b>	A statistical measure of the level of dispersion of a set of values around their mean; square root of the variance.
<b>True intake</b>	Actual intake, which cannot be observed in practice among free-living individuals.
<b>Usual intake</b>	Long-term average daily intake, taking into account both consumption and nonconsumption days.