

# measurement ERROR webinar series

## Assessing diet-health relationships with FFQ: focus on episodically-consumed dietary components (Webinar 8)

### Objectives:

- Review statistical risk models for evaluating diet-health relationships in nutritional epidemiology.
- Describe application of regression calibration to correct for FFQ measurement error using repeat short-term reference measurements in a substudy.
- With focus on episodically-consumed dietary components, describe application of a new methodology to carry out regression calibration in risk models with energy-adjusted dietary covariates.

### Recommended resources:

- Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. Measurement error in nonlinear models: a modern perspective, 2nd edition. Boca Raton, FL: Chapman and Hall CRC Press; 2006. Chapter 4, Regression calibration.
- Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. J Natl Cancer Inst. 2011;103(14):1086-92.
- Kipnis V, Freedman LS, Brown CC, Hartman A, Schatzkin A, Wacholder S. Interpretation of energy adjustment models for nutritional epidemiology. Am J Epidemiol. 1993;137:1376-80.
- Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, Subar AF, Tooze JA, Carroll RJ, Freedman LS. Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. Biometrics. 2009;65(4):1003-10.
- 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(9):1051-69; discussion 1071-3.
- Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr. 1997;65(4 Suppl):1220S-1228S; discussion 1229S-1231S.

### Key terms:

<b>Bivariate</b>	Having to do with two variables.
<b>Box-Cox transformation</b>	A type of power transformation; often applied to skewed data to lessen skewness or to approximate normality.
<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>Cohort study</b>	A study in which exposures of interest are assessed at baseline in a group (cohort) of people and health outcomes occurring over time (observed prospectively) are then related to baseline exposures.
<b>Consumption day</b>	A day on which a particular nutrient or food is consumed by a specific individual.
<b>Cox regression</b>	A statistical method for relating the time until a specified event (for example, a health outcome or mortality) to covariates of interest; also known as the proportional hazards model.
<b>Density model</b>	Regression model used for examining diet-health relationships in which nutrients or foods are expressed as densities (that is, ratios of nutrients or foods to energy).
<b>Energy adjustment</b>	Adjustment of nutrient intake for total energy intake.
<b>Epidemiology</b>	The study of the distribution and determinants of health outcomes or diseases among populations and the application of that study to enhancing public health.
<b>Episodically consumed dietary components</b>	Nutrients and foods that are not consumed on a daily basis by nearly everyone in the population and whose intake may therefore commonly be reported as zero on a particular day.
<b>Error-prone variable</b>	A variable typically measured with error.
<b>Exposure</b>	A potential determinant of a health or disease outcome; often a substance in the environment (for example, air pollution) or a personal habit (for example, dietary intake, smoking).
<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>Hazard ratio</b>	A ratio similar to relative risk; it expresses the relative effect of a variable on the risk of an event (such as the development of a disease) in the context of a Cox regression model.
<b>Latent variable</b>	A variable that is not directly observed but is inferred.
<b>Linear regression</b>	A statistical model that relates a dependent variable (for example, an outcome) to one or more independent variables (for example, exposures).
<b>Logistic regression</b>	Statistical model that relates a binary outcome to one or more independent variables, using the logit link.

<b>Main dietary instrument</b>	The primary dietary instrument used in a study, sometimes referred to as the study instrument; may be calibrated or validated using a reference instrument.
<b>Maximum likelihood estimation</b>	A technique used to estimate the parameters of statistical models, based on the principle that the best estimates of the parameters are those for which the observed data could most likely have arisen.
<b>Measurement error</b>	The difference between the observed or measured value and the true value.
<b>Multivariate</b>	Having to do with two or more variables.
<b>National Cancer Institute (NCI) method</b>	A unified approach for estimating usual intake distributions and predicting individual intakes for use in diet and health models; can be used for dietary components consumed nearly daily by nearly all persons and those consumed episodically.
<b>National Health and Nutrition Examination Survey (NHANES)</b>	A representative survey of the civilian, noninstitutionalized U.S. population conducted by the National Center for Health Statistics; used to monitor diet and study associations between diet, nutrition, and health.
<b>National Institutes of Health-American Association of Retired Persons Diet and Health Study (NIH-AARP)</b>	A diet and health cohort study of more than 500,000 Americans aged 50-71 years in 1995-1996.
<b>Never consumers</b>	Individuals who never consume a particular food or nutrient.
<b>Observing Protein and Energy Nutrition (OPEN)</b>	A study conducted by the National Cancer Institute in 1999-2000 to assess dietary measurement error using two self-report instruments (24HR and FFQ) and unbiased biomarkers of energy and protein intakes; included 484 men and women aged 40-69 years living in Montgomery County, Maryland.
<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>Person-specific random effect</b>	The difference between the within-person average value and the value predicted by covariates such as age and sex; both parts of the two-part National Cancer Institute method model include a person-specific random effect.
<b>Probit regression</b>	A statistical model for predicting the probability of a binary outcome using the probit link function.

<b>Prospective study</b>	A study in which participants are recruited and their exposures measured before the health outcome of interest has occurred.
<b>Random error</b>	A source of error that contributes variability (reduces precision) but does not influence the sample mean or median.
<b>Random within-person error</b>	Variation in the observed value of a variable when it is repeatedly measured in the same individual; for example, day-to-day variation in dietary intake reported using multiple 24-hour recalls.
<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>Reference dietary instrument</b>	An instrument that is administered in a substudy and is used to calibrate or validate the main or study instrument; examples include recovery biomarkers. The reference instrument is assumed to provide estimates that are closer to truth than the main instrument.
<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>Relative risk</b>	A statistical measure that quantifies the association between an exposure and a health outcome; often used in cohort studies.
<b>Short-term instrument</b>	A dietary instrument that captures intake over a short period of time, such as a food record or 24-hour recall.
<b>Simulation study</b>	A method used to validate statistical procedures that involves generating random samples from a hypothetical distribution and computing statistical estimates for each sample.
<b>Systematic error (bias)</b>	A source of error in which measurements consistently depart from the true value in the same direction; affects the sample mean or median and can result in incorrect estimates and conclusions.
<b>Transformation</b>	The application of a mathematical function (for example, the logarithm or the square root) to a set of values to create a new set of values.
<b>True intake</b>	Actual intake, which cannot be observed in practice among free-living individuals.

<b>Twenty-four-hour dietary recall (24HR)</b>	A dietary instrument that requires the respondent to remember and report all foods and beverages consumed in the preceding 24 hours or during the preceding day.
<b>Two-part model</b>	For the purposes of the webinar series, a statistical regression model developed by the National Cancer Institute for estimating usual intake of dietary components that are episodically consumed; models the probability of consuming the component on a particular day as well as the usual amount consumed on a consumption day.
<b>Unbiased instrument</b>	An instrument with only random error.
<b>Univariate</b>	Having to do with one variable.
<b>Usual amount consumed</b>	For episodically consumed dietary components, the usual amount consumed is the long-term average amount consumed on consumption days; when multiplied by the probability of consuming the dietary component, the product equals usual intake.
<b>Usual intake</b>	Long-term average daily intake, taking into account both consumption and nonconsumption days.