measurement ERROR webinar series

The problem of measurement error when examining diet-health relationships

(Webinar 6)

Objectives:

- Explain the types and magnitude of measurement error that occur in dietary data.
- Describe statistical models for evaluating diet-health relationships, including energy adjustment models.
- Describe the qualitative and quantitative impact of measurement error on studies of diet-health relationships.

Recommended resources:

- 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, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, Bingham S, Schoeller DA, Schatzkin A, Carroll RJ. Structure of dietary measurement error: results of the OPEN biomarker study. Am J Epidemiol. 2003;158(1):14-21; discussion 22-6.
- Schatzkin A, Subar AF, Moore S, Park Y, Potischman N, Thompson FE, Leitzmann M, Hollenbeck A, Morrissey KG, Kipnis V. Observational epidemiologic studies of nutrition and cancer: the next generation (with better observation). Cancer Epidemiol Biomarkers Prev. 2009;18(4):1026-32.
- Thiébaut AC, Kipnis V, Schatzkin A, Freedman LS. The role of dietary measurement error in investigating the hypothesized link between dietary fat intake and breast cancer – a story with twists and turns. Cancer Invest. 2008;26(1):68-73.
- 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:

Association	A relationship between two variables that is not necessarily causal.
Attenuation	Bias of the estimated regression coefficient in the direction of zero due to measurement error in a covariate; bias to the null.
Attenuation factor	The multiplicative factor by which an estimate of a regression coefficient is shrunk due to measurement error in a covariate.
Bias	Systematic deviation of observations or estimates from the truth.
Biomarker	For the purposes of the webinar series, a biological (usually biochemical) indicator or measure of dietary intake or nutritional status.

Case-control study	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.
Classical measurement error	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.
Cohort study	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.
Concentration biomarker	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.
Confidence interval	A range in which, for a specified degree of assurance, the true value of the parameter lies.
Confounding	Distortion of an association between an exposure and a health outcome by a third variable that is related to both.
Constant additive error	A component of systematic error that consists of a constant value that is added to the true value for each person.
Contamination factor	A value that indicates the magnitude of residual confounding in a regression model with multiple exposures measured with error.
Covariate	A variable that is related to the outcome or dependent variable in a regression model; may be referred to as an exposure.
Density model	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).
Energy adjustment	Adjustment of nutrient intake for total energy intake.
Explanatory variable	A variable thought to be related to an outcome in a regression model.
Flat slope	A syndrome affecting dietary data due to overreporting among those with the lowest levels of intake and underreporting among those with the highest levels of intake. When reported dietary intakes are regressed on true intakes, the result is a regression slope less than one.

Food frequency questionnaire (FFQ)	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.
Intake-related bias	Systematic deviation from the truth arising from correlation between error and true intake; for example, persons with low energy intake may overreport intake and persons with high energy intake may underreport intake.
Logistic regression	Statistical model that relates a binary outcome to one or more independent variables, using the logit link.
Measurement error	The difference between the observed or measured value and the true value.
Multivariate	Having to do with two or more variables.
Nutrient density	Ratio of nutrient intake to total energy intake, often expressed either as a percentage of total energy or amount per 1,000 kilocalories.
Observing Protein and Energy Nutrition (OPEN)	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.
Odds ratio	A statistical measure that quantifies the association between an exposure and a health outcome; often used in case-control studies.
Outcome	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.
Person-specific bias	The difference between an individual's reported intake averaged over many repeated measures and true usual intake, after taking constant additive error and intake-related bias into account. It is constant within an individual but randomly changes between individuals, with a mean of zero and constant variance.
Power	The probability that a test correctly rejects the null hypothesis when the alternative hypothesis is true.
Quantiles	Values that divide data or a distribution into equal-size groups; for example, quartiles are quantiles that divide the data into four equally sized groups.
Random error	A source of error that contributes variability (reduces precision) but does not influence the sample mean or median.

Random within-person error	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.
Recovery biomarker	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.
Regression model	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.
Relative risk	A statistical measure that quantifies the association between an exposure and a health outcome; often used in cohort studies.
Systematic error (bias)	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.
True intake	Actual intake, which cannot be observed in practice among free-living individuals.
Twenty-four-hour dietary recall (24HR)	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.
Unbiased instrument	An instrument with only random error.
Usual intake	Long-term average daily intake, taking into account both consumption and nonconsumption days.
Within-person variance	A measure of the variation in repeated observations of a variable in the same person. In dietary measurement using 24-hour recalls, it is the day-to-day variation in reported dietary intake of an individual.