

measurement ERROR webinar series

Assessing diet-health relationships using a short-term unbiased dietary instrument: focus on risk models with multiple dietary components (Webinar 12)

Objectives:

- Identify challenges in estimating relationships between a dietary exposure measured by repeated application of a short-term unbiased instrument and a health outcome in a risk model with multiple dietary components.
- Describe potential approaches to correct for within-person random measurement error in estimating relationships between a dietary exposure measured by a short-term unbiased instrument and a health outcome in a risk model with multiple dietary components.

Recommended resources:

- 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.
- Zhang S, Midthune D, Guenther PM, Krebs-Smith SM, Kipnis V, Dodd KW, Buckman DW, Tooze JA, Freedman L, Carroll RJ. A new multivariate measurement error model with zero-inflated dietary data, and its application to dietary assessment. *Ann Appl Stat*. 2011;5(2B):1456-87.

Key terms:

Box-Cox transformation	A type of power transformation; often applied to skewed data to lessen skewness or to approximate normality.
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.
Consumption day	A day on which a particular nutrient or food is consumed by a specific individual.
Cox regression	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.
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.
Epidemiology	The study of the distribution and determinants of health outcomes or diseases among populations and the application of that study to enhancing public health.
Episodically consumed dietary components	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.
Error-prone variable	A variable typically measured with error.
Exposure	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).
Hazard ratio	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.
Latent variable	A variable that is not directly observed but is inferred.
Linear regression	A statistical model that relates a dependent variable (for example, an outcome) to one or more independent variables (for example, exposures).
Logistic regression	Statistical model that relates a binary outcome to one or more independent variables, using the logit link.
Main dietary instrument	The primary dietary instrument used in a study, sometimes referred to as the study instrument; may be calibrated or validated using a reference instrument.
Markov chain Monte Carlo method	A technique used to estimate the parameters of statistical models through simulation.
Maximum likelihood estimation	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.
Measurement error	The difference between the observed or measured value and the true value.
Multivariate	Having to do with two or more variables.

National Cancer Institute (NCI) method	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.
National Health and Nutrition Examination Survey (NHANES)	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.
Nonepisodically consumed dietary components	A term describing nutrients and foods that are consumed nearly every day by nearly everyone in the population and whose intake may therefore rarely, if ever, be reported as zero on a particular day.
Odds ratio	A statistical measure that quantifies the association between an exposure and a health outcome; often used in case-control studies.
Person-specific random effect	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.
Probit regression	A statistical model for predicting the probability of a binary outcome using the probit link function.
Prospective study	A study in which participants are recruited and their exposures measured before the health outcome of interest has occurred.
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.
Regression calibration	A statistical method for correcting estimated regression coefficients for bias due to measurement error in one or more continuous covariates.
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.

Replicates/repeats	A repeated measure; for example, a second 24-hour recall or a second doubly labeled water measurement.
Short-term instrument	A dietary instrument that captures intake over a short period of time, such as a food record or 24-hour recall.
Simulation study	A method used to validate statistical procedures that involves generating random samples from a hypothetical distribution and computing statistical estimates for each sample.
Transformation	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.
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.
Two-part model	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.
Unbiased instrument	An instrument with only random error.
Usual amount consumed	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.
Usual intake	Long-term average daily intake, taking into account both consumption and nonconsumption days.