

# measurement ERROR webinar series

## Estimating usual intake distributions for dietary components consumed episodically (Webinar 3)

### Objectives:

- Define key concepts of food consumption related to usual intake estimation.
- Identify challenges in estimating usual intake for episodically-consumed dietary constituents.
- Explain statistical modeling approaches for estimating usual intake for episodically-consumed dietary constituents.
- Apply NCI macros.

### Recommended resources:

- Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, Tooze JA, Krebs-Smith SM. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *J Am Diet Assoc.* 2006;106(10):1640-50.
- Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, Guenther PM, Carroll RJ, Kipnis V. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc.* 2006;106(10):1575-87.

### Key terms:

<b>Back-transformation</b>	A mathematical technique used to restore a variable to its original scale after a transformation has been applied.
<b>Between-person variance</b>	A measure of the spread of values among persons.
<b>Bias</b>	Systematic deviation of observations or estimates from the truth.
<b>Box-Cox transformation</b>	A type of power transformation; often applied to skewed data to lessen skewness or to approximate normality.
<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>Consumption day</b>	A day on which a particular nutrient or food is consumed by a specific individual.
<b>Correlation</b>	A measure of linear association between two variables.
<b>Covariance</b>	A measure of how much two variables change in concert with each other.

<b>Covariate</b>	A variable that is related to the outcome or dependent variable in a regression model; may be referred to as an exposure.
<b>Day-of-week effect</b>	A phenomenon indicating how overall mean intake varies according to the day of the week.
<b>Dietary intake</b>	Intake from foods and beverages (excludes supplements).
<b>Dietary patterns</b>	For the purpose of this webinar series, this term refers to the combination of foods and beverages that constitute an individual's dietary intake over time.
<b>Distribution</b>	The pattern of values taken on by a random variable.
<b>Eating at America's Table Study (EATS)</b>	A study conducted by the National Cancer Institute in 1997-1999 to validate the Diet History Questionnaire, a food frequency questionnaire, and two dietary screeners; included a nationally representative sample of 1,640 men and women aged 20-70 years.
<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>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>Iowa State University (ISU) method</b>	A statistical modeling approach used to estimate distributions of usual intake.
<b>Logistic regression</b>	Statistical model that relates a binary outcome to one or more independent variables, using the logit link.
<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>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>Nonepisodically consumed dietary components</b>	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.

<b>One-part model</b>	For the purposes of this webinar series, this term refers to a model developed by the National Cancer Institute for estimating usual intake distributions of nonepisodically consumed dietary components and predicting individual intakes of nonepisodically consumed dietary components for use in diet-health models.
<b>Percentile</b>	The value below which the specified percent of observations fall.
<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>Probability</b>	The chance of a particular event or value; how likely an event is to occur.
<b>Probability distribution</b>	The pattern of values showing the relative frequencies associated with all possible values of a random variable in a population. Examples include Normal, t, F, Binomial, and Chi-square.
<b>Probit regression</b>	A statistical model for predicting the probability of a binary outcome using the probit link function.
<b>Quantiles</b>	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.
<b>Random error</b>	A source of error that contributes variability (reduces precision) but does not influence the sample mean or median.
<b>Random variable</b>	A characteristic of interest that takes on values not directly fixed by an experiment. Examples include dietary intake, height, and weight.
<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>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>Standard error</b>	The standard deviation of the sampling distribution of an estimated population parameter; used to assess the precision of an estimate.
<b>Surveillance</b>	A general term for monitoring; in the context of nutrition, refers to tracking the population's diet- and nutrition-related health events.
<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>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.
<b>Usual intake distribution</b>	A distribution that describes usual intakes, including the mean and percentiles, among a population.
<b>Variance</b>	A measure of the spread in a set of observations; it is equal to the mean squared difference between observations and their mean value.
<b>Within-person variance</b>	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.