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:


Key terms:

- **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.
- **Biomarker**: For the purposes of the webinar series, a biological (usually biochemical) indicator or measure of dietary intake or nutritional status.
- **Calibration equation**: 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.
**Calibration substudy**
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.

**Carotenoids in Age-Related Eye Disease Study (CAREDS)**
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).

**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.

**Causal/Causation**
A type of relationship between two variables in which a change in the value of one causes the value of the other to change.

**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.

**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.

**Confounding**
Distortion of an association between an exposure and a health outcome by a third variable that is related to both.

**Explanatory variable**
A variable thought to be related to an outcome in a regression model.

**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.

**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.
Mediation: 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.

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.

Power: The probability that a test correctly rejects the null hypothesis when the alternative hypothesis is true.

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 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.

Standard deviation: A statistical measure of the level of dispersion of a set of values around their mean; square root of the variance.

True intake: Actual intake, which cannot be observed in practice among free-living individuals.

Usual intake: Long-term average daily intake, taking into account both consumption and nonconsumption days.