Paper P6-4 ***Food Composition Data: Making Use of Variability***
Gustaaf Sevenhuysen, University of Manitoba, CANADA

FOOD COMPOSITION DATA: MAKING USE OF VARIABILITY

Sevenhuysen GP[1], Holcikova K[2], Holden J[3], and Kutka J[2].
[1] University of Manitoba -Canada
[2] Food Research Institute-Slovakia
[3] USDA - Washington DC

ABSTRACT

The mean value is the most common way to report the amount of a component in a food. In most databases the mean value is intended to reflect the average composition of similar food samples from a specified geographical area. However, the amounts of a food component in any one food may not be distributed normally, which can preclude the use of the mean, standard deviation and standard error. Though median values may better predict the content of other samples, such values can not be summed to calculate the content of a mixture of foods, or perform other linear operations. An alternative technique of presenting food composition data using percentile distributions is described, suitable for use with very small sample sizes. The technique allows users to determine the probability that a combination of food items contains levels of components recommended for nutrient intake or required for food product labeling. In addition, the technique can assist users to describe more precisely which characteristics of food items increase variability, such as growing conditions, storage or processing steps. Data from Eastern Europe and from The Carotenoid Project, USDA, is used to demonstrate the technique that provides probabilities of food component levels. The practical process of generating databank values, examples of data presentation and their use in food product development and dietetic practice are included.

PRESENTING FOOD COMPOSITION DATA

Problem

The amount of a nutrient or component in a food is most commonly expressed as the mean of several analytical estimates derived from different samples of the food. Standard deviations are used to show the extent to which actual food contents may vary from the mean. However, the calculation of standard deviations requires that the analytical values from a sample of food items are normally distributed. Since analytical values are frequently not normally distributed, standard deviations are often incorrect estimates of the probability with which given food component amounts occur.

Alternative measures of variability such as percentile values can represent variability in non-normal distributions of data. The 50^{th} percentile, or median value, is interpreted as the food component amount that the next food sample will contain with a 50% probability. However, percentile values obtained from several foods can not be added or averaged to find the percentiles for the recipe or daily intake the foods represent. As well, the majority of food composition data involves small sample sizes, e.g. 3 or 6 data values, for each food and nutrient combination, which prevents the use of traditional percentile calculations.

New technique

Create a percentile distribution for a food. Instead of single percentile values, group them according to the extent of variation due to the analytical technique used to generate the data. For example, the frequencies for each group of 4 percentiles can be used. If the technique is less precise, the frequencies of adjoining percentile groups would increase to show the increased probability. Create a percentile distribution for a mixture of foods by adding the nutrient amount values from different foods that are associated with the same percentile.

Create a point-to-point curve by connecting the frequency values of each bin, including all the ones that show zero frequencies. Divide the area under the curve in 100 equal parts by using vertical lines that intersect the horizontal scale and the curve. Record the value associated with areas of probability desired: 50%, 5%, 10%, 90% and 95%.

Application

The technique provides information about the likelihood that the amount of a nutrient or component in a food meets a given standard, because of the use of non-normal distributions and the use of percentiles. As a result, it is possible to quantify the chance that an individual's intake fails to meet the physiological requirement; or, to formulate a food product in order to be certain of minimum or maximum amounts of certain food components.