Food Chemistry 193 (2016) 141–147

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Validation of a sampling plan to generate food composition data

N.C. Sammán [a,*], M.A. Gimenez [a], N. Bassett [b], M.O. Lobo [a], M.E. Marcoleri [a]

[a] Departamento de Agroindustrias, Facultad de Ingeniería, Universidad Nacional de Jujuy, Avenida Italia esq. Martiarena, 4600 Jujuy, Argentina
[b] Instituto de Química Biológica, Facultad de Bioquímica, Química y Farmacia, Universidad Nacional de Tucumán – CONICET, Argentina

ARTICLE INFO

Article history:
Received 30 January 2014
Received in revised form 23 January 2015
Accepted 15 March 2015
Available online 4 April 2015

Keywords:
Sampling plan
Methodology validation
Food composition
Variability

ABSTRACT

A methodology to develop systematic plans for food sampling was proposed. Long life whole and skimmed milk, and sunflower oil were selected to validate the methodology in Argentina. Fatty acid profile in all foods, proximal composition, and calcium's content in milk were determined with AOAC methods. The number of samples (n) was calculated applying Cochran's formula with variation coefficients $\leqslant$12% and an estimate error (r) maximum permissible $\leqslant$5% for calcium content in milks and unsaturated fatty acids in oil. n were 9, 11 and 21 for long life whole and skimmed milk, and sunflower oil respectively. Sample units were randomly collected from production sites and sent to labs. Calculated r with experimental data was $\leqslant$10%, indicating high accuracy in the determination of analyte content of greater variability and reliability of the proposed sampling plan. The methodology is an adequate and useful tool to develop sampling plans for food composition analysis.

1. Introduction

Food composition data are widely used to monitor nutrient intake within populations, to develop national and regional nutrition policies and dietary guidelines, to plan and conduct health research related to dietary intake, and to support food trade. Composition data for foods must be appropriate in level of detail, documentation, statistical parameters and other information to support the needs of those organizations and investigators who will use the data for diverse applications. Today there is a demand for reliable food composition information that can be validated (Chu et al., 2009; Noncioli, 2013).

A food composition database (FCDB) is an electronic repository of food names and descriptions, food component values, and related documentation which are generated or acquired from various sources, including scientific literature. Whether the purpose is to generate data for a new database or to expand/upgrade the quality of existing data in a database, a well-constructed approach to sampling the key foods is required to ensure overall usefulness and quality of the database. The lists of foods that contribute in approximately the 80% of the intake of any one specific nutrient of the diet are identified as key foods (Haytowitz, Pehrsson, & Holden, 2000).

Food sampling methods are critical to generate reliable data that represent the composition of the food of interest. A sampling plan is a specific procedure for the selection, extraction, preservation, transportation and preparation of the parts to be removed from a population to serve as samples (Horwitz, 1990). The process of defining all types of units of food that constitute the population (i.e., foods) of interest is a primary step. A portion of a material selected from a larger quantity of material is called a sample (Greenfield & Southgate, 2003, chap 6; Holden & Davis, 1997, chap 12).

When a food is collected specifically to generate results for a FCDB it is of primary importance that samples are representative of the food item that is consumed or sold. Analysis of the nutritional composition of a food material depends on the successful completion of a number of different steps. The final report of the Third International Conference Food Data (1999), incorporated a list of the critical steps in the food sampling process. The list included: set up goals; foods to sample; food components to analyze; required number of samples; selection, preparation and transportation of samples; analytical procedure; statistical analysis and data reporting. The sampling plan also seeks to determine a representative average value for each analyte of interest, and to estimate variability of nutrients and foods.

After determining the ranked list of foods to be sampled (keyfoods) and the nutrients of interest, it will be important to determine where the food sample units will be selected. The systematic approach to determine where to sample will be called the sampling frame (Cochran, 1977), and must give information related to what, where, how much, and when to sample. The sampling frame must be actualized to secure reliable estimates with known

* Corresponding author. Tel.: +54 3884 493160; fax: +54 3884 493141.
E-mail address: nsamman@fi.unju.edu.ar (N.C. Sammán).

http://dx.doi.org/10.1016/j.foodchem.2015.03.083
0308-8146/

variability for the nutrient content of food and beverages consumed by the population (Pehrsson, Perry, & Daniel, 2013; Haytowitz, Pehrsson, & Holden, 2008).

To determine the required number of units (*n*), statistical levels of confidence must be set in order to reflect the variability in the nutrient food composition.

The design of a food sampling protocol requires a general knowledge of the food type or class, as well as information about the specific foods or products within that class (e.g., poultry products, parts, with/without skin). It is important to know the way(s) the food(s) are available (raw, refrigerated, frozen), prepared (fried, baked, cooked dish), distributed, and consumed. The decisions for foods, forms, nutrients, and sampling options should be driven by the sampling objective(s) (Haytowitz, Pehrsson, & Holden, 2000). The definition of the objectives will depend upon the resources available as well as the project goals to be achieved. The sampling plan to generate food composition data may have different goals. They can be applied to determine nationally representative estimates of the composition of foods (Galeazzi, Lima, Colugnati, Padovani, & Rodriguez-Amaya, 2002; Pehrsson, Haytowitz, Holden, Perry, & Beckler, 2000; Nickle & Pehrsson, 2013); to conduct a pilot study to determine the magnitude of variances associated with specific parameters (e.g., cultivar, feed, breed) (Davey et al., 2007); to generate comprehensive and representative data for a specific nutrient (Holden et al., 2005), among others.

Under the FAO project TCP/RLA/3107 "Developing Food Composition Data Base for Argentina, Chile and Paraguay to strengthen the international trade and consumers' protection", a probability-based methodology to systematically develop sampling plans for food analysis was proposed. The methodology discussed and approved in a Workshop of the FAO project, was included in the Sampling Workbook for Latin American Countries in process of publishing (Holden, Pehrsson, Perry, & Greenfield, 2012). It has a structure of five steps (Samman, Masson, de Pablo, & Ovelar, 2011) and was developed in order to ensure that the data are truly representative of the national food supply.

Usually a large number of samples, depending on the food type and variables affecting their composition, must be analyzed to obtain representative data (Galeazzi et al., 2002; Greenfield et al., 2009; Tarley, Visentainer, Matsushita, & de Souza, 2004). The objective of the proposed methodology is to select the main variables that influence the composition of each food, discarding the less important in order to reduce variability, and thereby the samples number analyzed, without losing its statistical significance.

Fig. 1 summarizes the main steps of the proposed methodology and Fig. 2 includes two approaches used to define where to sample according to production or consumption conditions. Food samples could be selected in shops, supermarkets, street vendors, farms or homes, or any other location in every region, or province in the country. Types of production or consumption of food may determine the decision of where to sample. In many countries of Latin America, including Argentina, statistics are unavailable on consumption data or sales distribution in supermarkets. It is therefore necessary to look for other alternative sources of information on where to sample.

If production is concentrated in a defined region or in few factories it would be convenient to take samples in manufacture/production places. In this case n could be distributed proportionally to production volume of each company. If production is widespread throughout the country, it would be desirable to take samples at points of consumption.

Cochran uses the human population as a model system. When addressing the development of the sampling frame, Cochran recommended that "*the population must be divided into parts that are called sampling units or simply units*". These units must cover

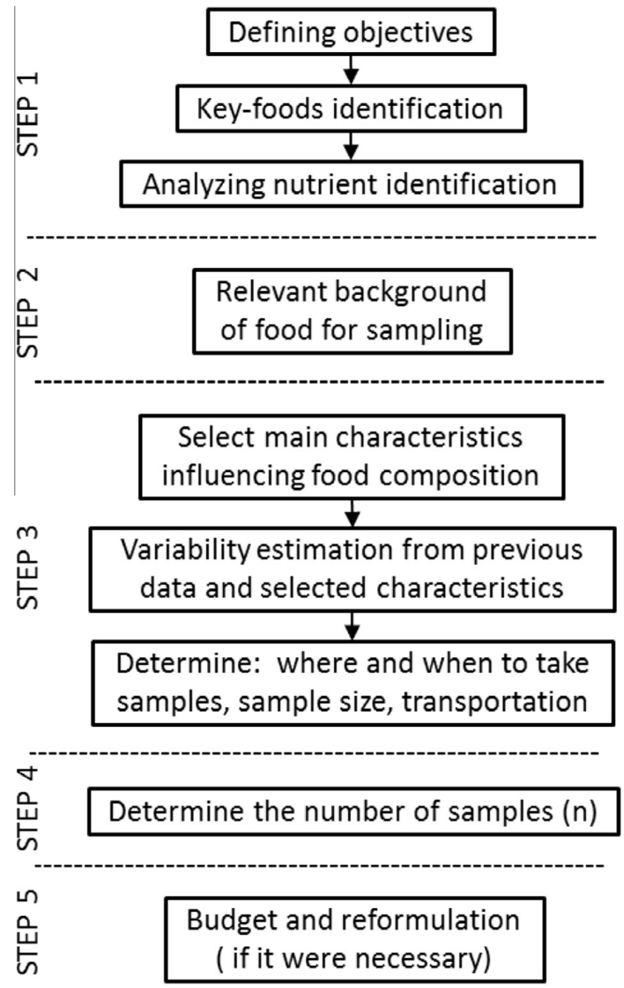

Fig. 1. Main steps of the proposed methodology to systematically develop sampling plans for food analysis in Latin America.

the whole population but they must not overlap, in the sense that every element in the population belongs to one and only one unit (Cochran, 1977). Therefore, it will be necessary to use a systematic approach to identify the list of all possible locations. From this list it will be necessary to select a subset of locations where the actual selection of the food sample units can take place. The number of locations could be determined, in part, by the objectives and scope of the study and in part, by the amount of funding and other resources available. In this case "*n*" is distributed randomly in selected cities with probability proportional to their population. For example, Perry, Pehrsson, and Holden (2003) obtained the U.S. National Census Data and identified the U.S. County as the unit of interest for the sampling frame of the National Food and Nutrient Analysis Program (NFNAP). The US Department of Agriculture's (USDA) Nutrient Data Laboratory develops a second revision of the NFNAP sampling plan implemented in 2012, for the national collection of food samples from retail outlets for nutrient analysis. In that case, Chromy's Procedure, a probability minimum replacement probability proportional to size sampling scheme was used (Pehrsson et al., 2013). In this way, in TACO project (Tabela Brasileira de Composicão de Alimentos), foods were selected from nine cities in the five official Brazilian geopolitical regions (Galeazzi et al., 2002).

Careful handling of food samples from the time of acquisition to the time of analysis is critical to ensure the integrity of the samples and subsequent generation of accurate nutrient values (Trainer et al., 2010; Westenbrink, Oseredczuk, Castanheira, & Roe, 2009).