Food and Agriculture Organization of the United Nations

# Principles about Sampling

Last update : January 2021

**U. Ruth Charrondiere, PhD**
**FAO**

Food and Agriculture Organization of the United Nations

# Sources of Variability

**Natural variation**

- Growing location
- Climate
- Breed/Cultivar
- Age/Ripeness
- Formulation/Recipe
- Brandnames

**Other variations**

- Analytical methods
- Nutrient definition including conversion factors

Food and Agriculture Organization
of the United Nations

# Nutrient composition of banana cultivars

| Banana Variety | Edible Portion | Water g | Energy kJ (kcal) | Calcium mg | Phos mg | Iron mg | ß carotene mcg |
|---|---|---|---|---|---|---|---|
| Cavendish | 64 | 74.4 | 435 (104) | 139 | 20 | 0.8 | 75 |
| Botoan | 57 | 74.4 | 422 (101) | 21 | 27 | 0.4 | 25 |
| Ternatensis | 62 | 66.3 | 552 (132) | 15 | 19 | 0.9 | 370 |
| Lacatan | 69 | 68 | 527 (126) | 21 | 34 | 0.8 | 360 |
| Violacea | 67 | 73.1 | 447 (107) | 19 | 21 | 0.7 | 285 |
| Compressa | 57 | 72.2 | 460 (110) | 23 | 36 | 0.9 | 190 |
| Ternatensis | 64 | 66.2 | 560 (134) | 11 | 24 | 0.7 | 325 |
| Tuldoc | 76 | 74.8 | 414 (99) | 26 | 28 | 1.6 | 1370 |
| *Uht en yap* | | 69.5 | | | | | 2780 |

Philippine Food composition tables, 1997 and Englberger et al. 2003 JFCA

# Carotene content of mango varieties

| Mango variety and carotene | Immature (mcg/100g pulp) | Partially mature (mcg/100g pulp) | Mature (mcg/100g pulp) |
|---|---|---|---|
| Mango Badami | | | |
| - beta carotene | 20 | 1130 | 4520 |
| - gamma carotene | trace | 10 | 20 |
| - beta cryptoxanthin | trace | 50 | 40 |
| Mango Keitt | | | |
| - beta carotene | 170 | 420 | 670 |
| - beta cryptoxanthin | trace | trace | 20 |
| Mango Tommy Atkins | | | |
| - beta carotene | 200 | 400 | 580 |
| - beta cryptoxanthin | 10 | 10 | 30 |

Adapted from Rodriguez-Amaya (1987)

Food and Agriculture Organization of the United Nations

# Basic terms (1)

Table 5.1 Definition of terms used in sampling of food for a nutritional database

| *Term* | *Definition* | *Comments on application in food composition studies* |
|---|---|---|
| Sample | A portion selected from a larger quantity of material | A general term for a unit taken from the total amount (the population) of a food |
| Sampling protocol | A predetermined procedure for the selection, withdrawal, preservation and preparation of the sample | Sometimes called a sampling plan |
| Characteristic | The property or constituent that is to be measured or noted | Description of the food, nutrient and other analyses |
| Homogeneity | The extent to which a property or constituent is uniformly distributed | Foods are usually heterogeneous or must be assumed to be so |
| Sampling error | The part of the total error associated with using only a fraction of the total population of food and extrapolating it to the whole population. This arises from the heterogeneity of the population | Because of the heterogeneous nature of foods, replicate samples must always be taken when estimating the composition of the population of a food |
| Batch | A quantity of food that is known, or assumed, to be produced under uniform conditions | Batch numbers should always be noted when sampling foods |
| Unit | Each of the discrete, identifiable units of food that are suitable for removal from the population as samples and that can be individually described, analysed or combined | These units form the basis of most food analysis work (e.g. an apple, a bunch of bananas, a can of beans, a prepared dish) |

Source: Greenfield & Southgate, 2003

# Basic terms (2)

**Table 5.3** Main sampling methods used in nutrient composition studies

| *Method* | *Definition and characteristics* | *Notes on application* |
|---|---|---|
| Random sampling | Samples are taken in a way that ensures that any one unit has an equal chance of being included | The theoretical ideal but rarely practicable when sampling foods for nutritional databases |
| Stratified sampling | Units of sampling are taken from defined strata (subparts) of the parent population. Within each stratum the samples are taken randomly | Often the most suitable method for use in database work. Strata may be regional, seasonal, retail sale point, etc., as defined by knowledge of the food being studied |
| Selective sampling | Samples are taken according to a sampling plan that excludes material with certain characteristics or selects only those with defined characteristics | Most commonly used in the analysis of contaminants. Can be used, with caution, for database work |
| Convenience sampling | Samples are taken on the basis of accessibility, expediency, cost or other reason not directly concerned with sampling parameters | Rarely suitable for database work but may be the only practicable way to sample wild or uncultivated foods or composite dishes from households |

Source: Greenfield & Southgate, 2003

Food and Agriculture Organization of the United Nations

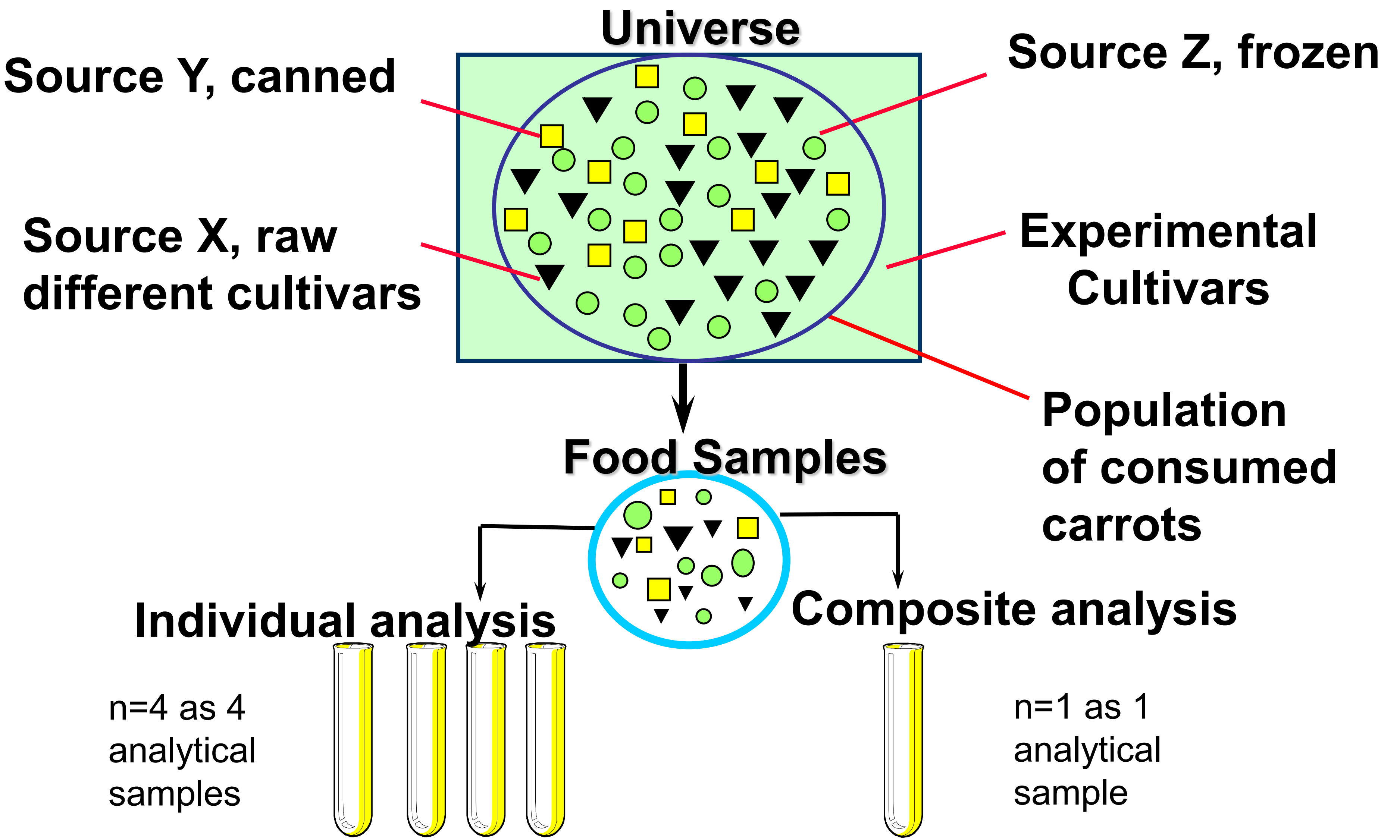
# Basic terms (3)

Table 5.4 Summary of stages in sampling and preparation of samples in food composition studies

| *Terms* | *Description* | *Main use in food composition studies* |
|---|---|---|
| Primary sample | The collection of one or more units initially taken from the total population of the food | The usual starting point in compositional studies. The ideal is the collection of several replicates that are treated separately. Primary samples are often mixed to form composites |
| Reduced sample | A representative part of the primary sample obtained by a division or reduction process | Frequently used to reduce the primary sample to a more manageable weight |
| Composite sample | Mixtures formed by combining primary samples | Frequently used in food composition studies. Composites may be samples of the same food or combinations of different brands or cultivars |
| Laboratory sample | The sample sent to or received by the laboratory | The primary sample (or a reduced sample) often requires further handling in the laboratory (e.g. thawing, cooking, separation of inedible matter). The edible portion may need further reduction or mixing |
| Analytical sample | The portion prepared from the laboratory sample from which the portions for analysis are taken | This is usually the form in which the food samples are prepared for analysis |
| Analytical portion | The quantity of food of the proper weight for each analytical measurement | The analysis of duplicate analytical portions is the minimum acceptable; several replications are preferable |

Source: Greenfield & Southgate, 2003

# Carrots




Source: adapted from Joanne Holden

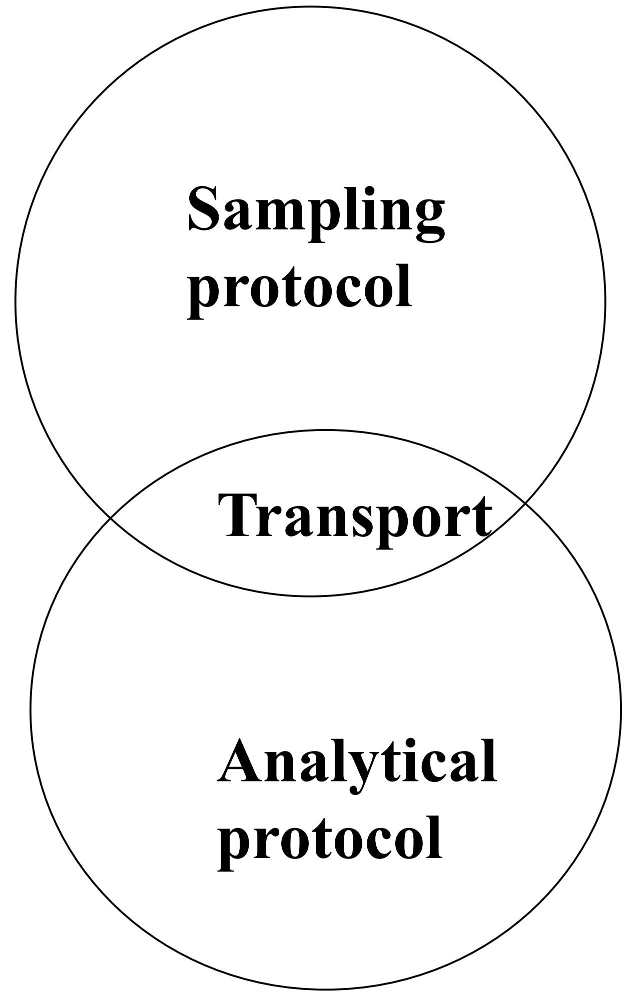
# What can we achieve?

- ➔ with individual samples mean + variability
- ➔ with composite sample only mean but no variability between samples

- ➔ the larger number of units the better the estimation of the true mean

# Sampling and analytical protocol

To be elaborated together by compilers and analysts to ensure:
- a better data quality
- a smooth process from sample collection to analysis
- grouped collection and analysis of similar foods (advantage for sample preparation and analysis, e.g. calibration)



# Responsibility for preparing the combined protocols (1)

- Database compilers
  - control the sampling and analytical work
  - responsible, in collaboration with the analysts, for preparing the written combined protocols
- Analysts
  - assure that the plan is feasible in terms of transport to the laboratories, their capacities to store, handle and analyze the samples adequately

# Responsibility for preparing the combined protocols (2)

**In most countries, sampling and analytical work may be contracted**

- then the compilers' input may be restricted to establishing the broad outlines of the work required
- the written contract should include the following:
  - the preferred analytical methods per nutrient
  - the minimum quality assurance
  - the components to be analyzed with their definition, units (preferably metric units and precision of denominator, e.g. if total food or edible portion), and an internationally recognized component identifier, e.g. INFOODS tagnames
  - the foods to be analyzed
  - storage (temperature, duration) and sample handling
  - date to submit results

Food and Agriculture Organization
of the United Nations

# Sampling plan

- Is based on a clear understanding of the population of foods studied
- Describes the sampling type and its particular design
- Defines the sampling sites
- Provides a detailed description of the sampling process to be undertaken
- Is written with the objective of ensuring that no changes in composition occur between collection and analysis
- Defines the number and size of samples

# Sampling strategy

- Population based
  - identify where population is located who eats the foods
  - mostly sampling in big cities
- Based on agricultural production
  - identify where the food is produced
  - sampling in production area

- → mostly population based and mostly stratified sampling

Food and Agriculture Organization of the United Nations

# Knowledge of Foods

- Where and when the food is consumed/produced?
- What forms/types/brands of the food are consumed?
- Where is the food produced?  When?
- By whom consumed? Specific groups, urban/rural
- How is the food prepared?
- Is the food prepared from a recipe or formulation?
- Statistics?

# Sources of Information about Foods

- Government/Industry reports
- Census Data
- Production/Sales Data
- Expert Knowledge

# Number of samples

**Sample size ≥ ($t$ $\alpha$ $n$-1)$^2$ SD$^2$/(accuracy x mean) $^2$**

**<u>Examples of value for *t*</u>**

- sample size 10. For $\alpha$ = 0.05 and df = 9, $t$ for 0.025 ($\alpha/2$) = 2.262. ➔ $t^2$ = 5.1166
- For a sample size of 20, and $\alpha$ = 0.05, $df$ = 19
- $t$ for 0.025 ($\alpha/2$) = 2.093. Thus $t^2$ = 4.3806

# Number of samples

| Parameter | Moisture (g/100g) | Fat (g/100g) | Cholesterol (mg/100g) |
|---|---|---|---|
| Actual sample size | 10 | 10 | 10 |
| Actual mean | 49.9 | 13.4 | 16 |
| Actual standard deviation (SD) | 8.5 | 3.9 | 6.7 |
| $SD^2$ | 72.25 | 15.21 | 44.89 |
| $t_{(\alpha = 0.1)}$ ($t_{(\alpha = 0.05)}$) | 1.860 (2.262) | 1.860 (2.262) | 1.860 (2.262) |
| $t^2_{(\alpha = 0.1)}$ ($t^2_{(\alpha = 0.05)}$) | 3.4596 (5.1166) | 3.4596 (5.1166) | 3.4596 (5.1166) |
| $t^2 \times SD^2$ | 249.96 (369.67) | 52.62 (77.82) | 155.30 (229.68) |
| Accuracy set at | 0.1 (0.05) | 0.1 (0.05) | 0.1 (0.05) |
| Accuracy x mean | 4.99 (2.495) | 1.34 (0.67) | 1.6 (0.8) |
| $(\text{Accuracy x mean})^2$ | 24.9 (6.225) | 1.7956 (0.4489) | 2.56 (0.64) |
| Sample size required for accuracy = 0.1 | 249.96/24.9 = 10 | 52.62 /1.7956 = 29 | 155.30 /2.56 = 60 |
| Sample size required for accuracy = 0.05 | 369.67/6.225 = 59 | 77.82/0.4489 = 173 | 229.68/0.64 = 359 |

Source: Greenfield & Southgate, 2003

# Size and number of samples

Application of equation requires knowledge on some parameters (mean, SD) of the food to be acquired through:

- pilot analytical studies (best option)
- literature
- intuitive guesses (least best option)

# Minimal number of samples

- most sampling schemes adopt a standard of at least 10 foods sampled and 3 analytical samples
- US nutrition labelling requires 12

# Sample collection

- collectors need exact written instructions what to collect, how much and where (with alternative if exact sample is not available) and how to send it when to the laboratory
- collectors need training
- each sample needs to be coded and documented before being sent as quickly as possible to laboratory

# Special considerations for transportation

Main objective: to conserve all nutrients to represent levels in food as consumed and avoid contamination:

- maintain water (avoid drying out) and vitamin content
- use sterile containers, preferably sealed
- if frozen, sample needs to be kept frozen (choose appropriate container and travel time – if time too long to laboratory to assure frozen state select other good quality laboratory)

Food and Agriculture Organization of the United Nations

# Sampling preparations, handling and storage

Main objective: to conserve all nutrients as initially present in the food while avoiding contamination

- foods are handled differently depending on form, size, matrix and nutrient to be analysed
- storage capacity and possibility determine when food has to be analyzed (e.g. if no deep-freezing possible - normally below 70 degree Celsius - food needs to be analyzed quickly except for minerals if the sample is dried)

# Example: NFNAP Sampling in for USDA database

- 12 or 24 locations (national): 4 regions (according to population distribution) with 3 strata each, and every strata 1 metropolitan statistical area with 1 urban and 1 rural location. 1 outlet in each location
- brands selected proportional to market share (1% or more of food consumed)
- → covering seasonal and regional differences, including urban vs. rural, and specific populations e.g. tribes
- → purpose: monitoring of Key Foods

Source: Joanne Holden

Food and Agriculture Organization of the United Nations

# Sampling for food biodiversity

- Decide for most important features to be included in sampling scheme
- Sample for varieties
- Generate analytical data for varieties
- Compile these data systematically and centrally
- Collect diet survey data on varieties
- Consider nutrient content in crop variety promotions

Food and Agriculture Organization of the United Nations

# Conclusions about sampling

- Without adequate sampling the analytical value cannot represent the food as consumed in the population throughout the year
- Sample for varieties, for regional and seasonal variations, different brands and foods of specific populations
- Generate also analytical data for varieties
- Avoid sampling errors

**For more on food composition, visit www.fao.org/infoods**