INTERNATIONAL CONFERENCE 2ND IMEKOFOODS

RECIPE CALCULATION:
HOW TO HANDLE VARIABILITY AND UNCERTAINTY?

Nadia Bastide[1], Delphine Lioger[2], Francisco Deolarte[3], Hervé This[4]

[1] Innit SAS, Inra-AgroParisTech International Centre for Molecular Gastronomy, Paris, France, nadia.bastide@agroparistech.fr; [2] Innit SAS, Paris, France, delphine.lioger@innit.com; [3] Innit Inc, Redwood city, California, United States, francisco.deolarte@innit.com; [4] UMR GENIAL, Université Paris-Saclay, Massy, France; Inra-AgroParisTech International Centre for Molecular Gastronomy, Paris, France, herve.this@agroparistech.fr

Abstract − One of the main limit of evaluating the nutrient content of mixed dishes is variability and uncertainty of data. Our aim is to propose a method for recipe calculation taking in account this issue. We estimate here final uncertainty and variability of nutrient values in recipe according to initial choices of data. This method allows to take in account a large range of initial data as long as there is no outlier. This estimation may be a valued tool when experimental data are not available.

Keywords: Recipe calculation, variability, uncertainty, databases.

1. INTRODUCTION

Most item in diet worldwide are multi-ingredient foods, i.e. contains more than one single ingredient [1]. The knowledge of the composition of these foods is essential for scientists and practitioners working in the field of nutrition, for instance to plan the dietary adequacy of meals and diets, to assess dietary status of population groups, and to explore relationship between diet and chronic diseases in epidemiological studies [2].

Direct analysis of foods is the preferred method to obtain nutrient composition data for foods of interest. But analytic data do not exist for most multi-ingredient foods and recipe analysis is a cost-effective alternative to estimate the nutrient content of foods [1].

Recipe calculation presents a number of inherent difficulties, mainly because of the abundance and diversity of foods and sources [3]. Indeed, in order to calculate the nutrient content of a dish, the following steps are needed: establish the weight of each ingredient (edible weight, converted to grams); assign nutrient data to each ingredient (a suitable nutrient profile is needed, it has to be adapted to the country); estimate the weight change during thermal processing (cooking) (the yield factors); and estimate vitamins and nutrient losses during thermal processing (the retention factor) [3]. There are several methods to calculate recipes. The most recommended one is the "mixed method", which consists in applying the yield factor at the recipe level, and the retention factor at the ingredient level [3][4].

Whatever the method used, the main challenge is to find appropriate values for each ingredient at each step of the calculation. There is indeed variability in the composition of food ingredients (quantities, country of origin, species), and uncertainty in nutritional data, yield and retention factors (measurement methods, sources) [5]. To deal with this issue, one has first to make choices, and to eliminate outliers and non-referenced values. But how to manage remaining variability and uncertainty? To our knowledge, there is today no reported method in food composition data allowing to take it into account this issue in recipe calculation.

In sciences of nature, this issue had already been raised with the problem of repeated measurements, and uncertainty of measuring devices. The solution adopted is calculation of uncertainty with the use of differential calculus, and there is a "guide to the expression of uncertainty in measurement" from the Joint Committee for Guides in Metrology (JCGM), which establishes general rules for evaluating and expression the uncertainty in measurement(s) which can be used in many fields [6]. This guide is also applicable to "evaluating and expression of the uncertainty associated with the conceptual design and theoretical analysis of experiments, methods of measurement, and complex components and

systems".

Our aim in the present study is to apply lessons from sciences of nature to food composition analysis by applying JCGM rules to the field of recipe calculation, for an estimation of uncertainty of nutrients values allowing more precise recipe analysis.

2. MATERIAL AND METHODS

2.1. Sources of data

Recipes were taken from popular US websites of recipes such as Allrecipe, Good Houseekeeping and NYT cooking. We chose recipes representative of different type of dishes: White cooked meat - roasted vegetable chicken (Allrecipe, [7]); Red cooked meat - slow cooker Indian lamb (Good Houseekeeping, [8]); Cooked fish - monkfish with mashed potatoes and thyme (NYT Cooking, [9]); Cooked vegetable - roasted broccoli rabe (NYT Cooking, [10]); Cooked dessert - Free-Form Apple tart (NYT Cooking, [11]).

Nutritional data were taken in the last release of USDA database, sr28 [12]. Weight of ingredients were taken from the recipe when given (with waste), otherwise in USDA database (edible portion without refuse). The edible factor was applied in calculation in the first case only.

Retention factors were taken from the 2008 EuroFIR report [13]. Yield factors were taken from USDA and Bognar databases [14][15].

2.2. Recipe calculation

For recipe calculation, we used the mixed method as recommended by FAO and EuroFIR [3][4]. Retention factors were applied at the ingredient level, and yield factors at the recipe level. Data used in calculation were:

- Number of items of ingredient *i* in total recipe *R*: $n[R, i]$
- Mass of each item of ingredient *i* (g): $m[$
- Edible portion of each item of ingredient (%): $e[R,i]$ (imputed to 1 when $m[R,i]$ directly the edible portion, taken from USDA database).
- Nutrient value of nutrient x per 100g edible portion of ingredient *i* in recipe *R*: $NV[R, i, x]$
- Yield factor of total recipe *R*: $YF[R]$
- Retention factor of nutrient *x* for ingredient *i* in recipe *R*: $RF[R, i, x]$

The different steps of calculation were the following:

1. Total mass of raw recipe: Mr[R]

$$Mr[R] = \sum_{i=1}^{Ni} (n[R, i] \cdot m[R,i] \cdot e[R,i]) \quad (1)$$

2. Total mass of cooked recipe: Mc[R]

$$Mc[R] = YF[R] \cdot Mr[R] \quad (2)$$

3. Nutrient value in total cooked recipe for nutrient x: NVc[R,x]

$$NVc[R, x]$$
$$= \sum_{i=1}^{Ni} \frac{NV[R, i, x] \cdot M[R, i] \cdot RF[R, i, x]}{100} \quad (3)$$

4. Nutrient value per 100 g edible portion of cooked recipe for nutrient x: NVcs[R,x]

$$NVcs[R, x] = \frac{NVc[R, x] \cdot 100}{Mc[R]} \quad (4)$$

An Access database was made in order to gather pertinent values (Microsoft office, Access 2016). All calculations were made with Maple 18 (Maplesoft, Waterloo Maple Inc, Ontario, Canada).

2.3. Uncertainty and variability calculation

For uncertainty and variability calculation, we used the formula adapted from differential calculus by the Joint Committee for Guides in Metrology for the estimation of the uncertainty of measurements [6].

The principle is the following: for a differentiable function *f*, the error due to variations *Δx[i]* of the variable *x[i]* is:

$$\Delta f = \sqrt{\sum_{j=1}^{p} \left(\left(\frac{\partial f}{\partial x[j]} \cdot \Delta x[j] \right)^2 \right)} \quad (5)$$

3. RESULTS AND DISCUSSION

3.1. Uncertainty and variability formula adapted to recipe calculation

The formula (5) was applied to equation (1) to (4), with *Δn[R, i]*, *Δm[R, i]*, *Δe[R, i]*, *ΔNV[R, i, x]*, *ΔYF[R]* and *ΔRF[R, i, x]* equal to standard deviation

of values if several values, to standard deviation provided in database if one value, and to 0 in all other cases. The resulting formula are the following:

$$\Delta Mr[R] = \sqrt{\sum_{i=1}^{Ni} \begin{pmatrix} (\Delta n[R, i] \cdot m[R,i] \cdot e[R,i])^2 \\ +(n[R, i] \cdot \Delta m[R,i] \cdot e[R,i])^2 \\ +(n[R, i] \cdot m[R,i] \cdot \Delta e[R,i])^2 \end{pmatrix}} \quad (6)$$

$$\Delta Mc[R] = \sqrt{\begin{array}{l}(YF[R] \cdot \Delta Mr[R])^2 \\ +(\Delta YF[R] \cdot Mr[R])^2\end{array}} \quad (7)$$

$$\Delta NVc[R,x] = \sqrt{\begin{array}{l}\left(\dfrac{\partial NVc[R,x]}{\partial NV[R,x,i]} \cdot \Delta NV[R, x, i]\right)^2 \\[2mm] +\left(\dfrac{\partial NVc[R,x]}{\partial n[R,i]} \cdot \Delta n[R, i]\right)^2 \\[2mm] +\left(\dfrac{\partial NVc[R,x]}{\partial e[R,i]} \cdot \Delta e[R,i]\right)^2 \\[2mm] +\left(\dfrac{\partial NVc[R,x]}{\partial m[R,i]} \cdot \Delta m[R,i]\right)^2 \\[2mm] +\left(\dfrac{\partial NVc[R,x]}{\partial RF[R,x,i]} \cdot \Delta RF[R,x,i]\right)^2\end{array}} \quad (8)$$

$$\Delta NVcs[R,x] = 100 \cdot \sqrt{\begin{array}{l}\left(\dfrac{1}{Mc[R]} \cdot \Delta NVc[R, x]\right)^2 \\[2mm] +\left(\dfrac{NVc[R,x]}{Mc[R]^2} \cdot \Delta Mc[R]\right)^2\end{array}} \quad (9)$$

3.2. Practical application to recipe calculation

There are several options to calculate recipes. Among data entered in calculation, data which depends the most of the food compiler are:

- The number of items of ingredient i in total recipe R: n[R, i];
- The mass of each item of ingredient i (g): m[R,i];
- The nutrient value of nutrient x per 100g edible portion of ingredient i in recipe R: NV[R, i, x], according to the choice of food item by the food compiler
- The yield factor of total recipe R: YF[R]

Examples of recipes were calculated several times with several possible options, in order to see the variation in nutrient content per 100g edible portion cooked recipe NVcs[R, x] according to these choices.

To illustrate the role of number of ingredients on final nutrient value, we have chosen an example of

recipe, "Free Form Apple Tart" [11], which contains flour, sugar, butter, egg yolk, cream, and apples. In the recipe, it is specified "3 or 4 medium apples" Table 1 shows variations of mains nutrients per 100g edible portion according to the number of apples entered in recipe. In a logical manner, the more there are apples, the less there are macronutrients and energy per 100g of recipe. Inversely, the more there are apple, the more there are nutrients like fructose, glucose or vitamin C, abundant in apples (table 1).

Table 1: Nutrients values per 100g edible portion (±uncertainty calculated with formula (9)) of recipe "Free Form Apple Tart" [11] according to number of apple entered in recipe.

Number of apples in recipe	3 medium apples	3.5 medium apples	4 medium apples
Energy (kcal)	251.4±17.2	235.7±32.6	222.3±17.4
Carbohydrates (g)	26.0±2.0	25.1±3.9	24.3±2.2
Total lipids (g)	15.6±1.1	14.4±1.9	13.3±1.0
Protein (g)	2.9±0.2	2.7±0.4	2.5±0.2
Fructose (g)	3.7±0.6	4.0±1.1	4.2±0.7
Glucose (g)	2.0±0.4	2.1±0.7	2.3±0.5
Vitamin C (mg)	2.0±0.3	2.1±0.6	2.2±0.3

Table 1 shows that the choice of food compiler has an impact on final recipe composition. Nevertheless, choosing 3.5 apples, with uncertainty calculation allows to have a medium result, taking in account all possible variations, like presented in figure 1, with the example of energy value.
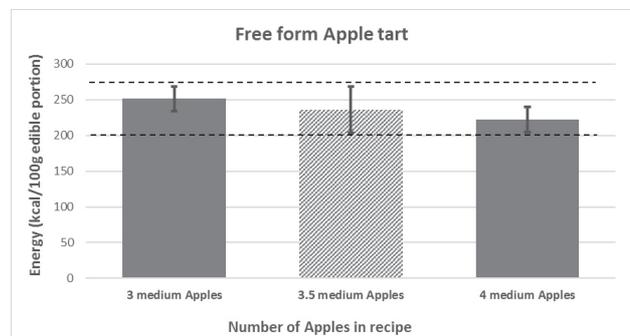

Fig. 1: Energy value per 100g edible portion (±uncertainty calculated with formula (9)) of recipe "Free Form Apple Tart" [11] according to number of apple entered in recipe.

The problematic is quite similar with weight of ingredients. In a recipe of roasted chicken with vegetables [7], there are carrots, celery, onions and potatoes. The food compiler can either choose the medium, large or small size of each item from USDA database, or mean of the three. This choice has an impact on nutrient final composition of the recipe per 100g, and using mean of small, medium and large sizes allows to have a medium value taking in account all possible variations (see Fig.2).

Choice of nutrients values present more difficulties. Indeed, in many times there are several possibilities in USDA database for one ingredient in recipe. For instance, in previous recipe of roasted vegetable chicken, the type of potato is not specified in recipe. From USDA database, food compiler can choose either: 11354 *"potatoes, white, flesh and skin, raw";* 11353 *"potatoes, russet, flesh and skin, raw"*; 11355 *"potatoes, red, flesh and skin, raw";* 11352 *"potatoes, flesh and skin, raw"*. Some possible options are the following:

- **Option 1**: Choosing medium size for all ingredients with item white potatoes (11354), the most used.
- **Option 2**: Choosing mean of all sizes for each ingredient with mean of items white, russet, and red potato (11353, 11354, 11355) for more accuracy.
- **Option 3**: Choosing mean of all sizes for each ingredient and item "potatoes, non-specified" in USDA (11352), which seemed to gather all others potatoes.

Fig. 2-4 shows the mass of total recipe, the proximate and vitamin values according to the option chosen:
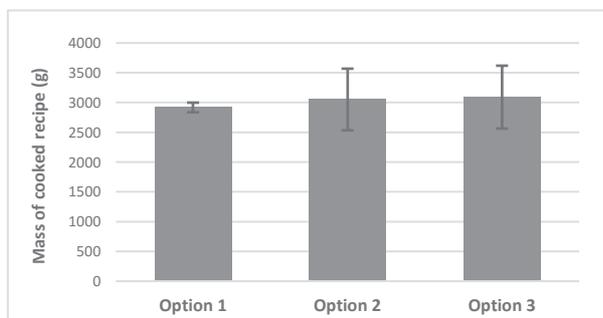

Fig 2: Mass of cooked recipe (±uncertainty calculated with formula (7)) according to options 1-3 in "Roasted Vegetables Chicken" recipe [7].

When the medium size for all ingredients is considered (option 1), the mass of cooked recipe is around 100 g less than when medium of all sizes for each ingredient is chosen.
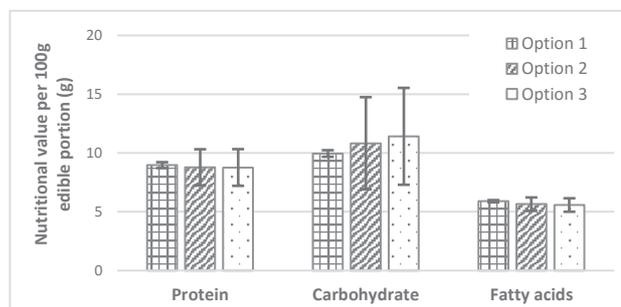

Fig 3: Proximate values per 100g edible portion (±uncertainty calculated with formula (9)) according to option 1-3 in "Roasted Vegetables Chicken" recipe [7]
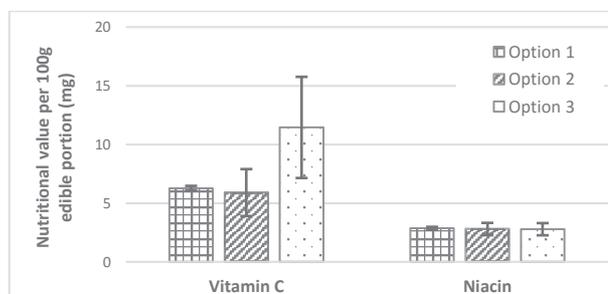

Fig 4: Vitamin values per 100g edible portion (±uncertainty calculated with formula (9)) according to option 1-3 in "Roasted Vegetables Chicken" recipe [7]

While there are no differences for proximate and niacin values, the vitamin C value is multiplied by a factor two between option 1 and 2 and option 3, with a huge uncertainty (35%). A careful analysis of data and their sources has shown that data for vitamin C in item 11352 was very different from the three others, and non-referenced (Table 2). Value of vitamin C for item 11352 is an outlier which should be removed before calculation.

Table 2: Nutrients values per 100g edible portion potatoes raw in USDA database [12].

USDA item	Vitamin C/100g edible portion	source
11352	**19.7 mg**	No source
11353	5.7 mg	National Food and Nutrient Analysis program
11354	9.1 mg	
11355	8.6 mg	

An additional issue regarding nutritional values is standard deviation given by USDA directly [12]. They are taken in account in uncertainty calculation and can induce until 26% of uncertainty for vitamin C in a recipe of roasted broccoli rabe [10], where it is the only parameter to change (table 4).

Table 4: Nutrients values per 100g edible portion (±uncertainty calculated with formula (9)) of recipe "Roasted Broccoli Rabe" [10].

Roasted Broccoli Rabe	Nutrient value/100 g edible portion	Uncertainty	% uncertainty
Vitamin C (mg)	16.4	4.2	26%
Iron (mg)	2.2	0.4	16%
Potassium (mg)	201	30	15%
Phosphorus (mg)	74.4	9	12%
Protein (g)	3.2	0.3	10%
Total lipid (g)	6.4	0.05	1%
Carbohydrates (g)	3.3	0	0%

The last point is the choice of the yield factor: there are two main sources, Bognar 2002 and USDA 2012 [14][15]. USDA was recently updated but gives only yield factors for meat. Bognar is more complete, purposing yield factor for "edible part" and "with waste" for most of existing foods. For instance, for a recipe of lamb stewed [8]: USDA yield factor for *"Lamb, retail cuts, shoulder, shank, breast, flank"* is 0.6. Bognar yield factor for *"Pot roast lamb, mutton (leg) small piece (20-500 g, high < 2.5 cm), with bones and gravy, meat and gravy part"* is 0.94 "edible part", and 1.06 "with waste". Fig.5 compare total weight of recipe according to different options. Like lamb is "cut for stew" in the recipe, we only have edible part. Fig. 4 shows the weight of cooked recipe according to different choices of yield factors.
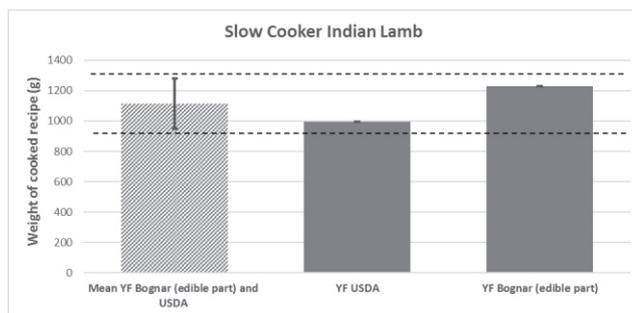

Fig. 5: Weight (±uncertainty calculated with formula (7)) of total recipe "slow cooker Indian lamb" [8] according to choice of yield factor.

We can see here that the choice of the yield factor can change weight of total recipe of nearly 20%. Nevertheless, with uncertainty calculation, we have a medium value more or less 15%, which allow to take in account all possible values.

The previous cases were examples in which choices from the food compiler were involved and induced uncertainty. In some cases, easier, there are very few choices to do, and uncertainty is very low, like for instance, for a monkfish with mashed potatoes recipe [9]. In this recipe number, weight and species of ingredients are specified, and there is no existing yield factor in USDA for fish. In consequence the only remaining uncertainty is due to standard deviation given by USDA in its nutritional table. In this case, uncertainty is below 10% for nearly all nutrients and below 1% for more than half of the nutrients.

3.3. Learning from this practical application

Food compilers have to face choices: number and weight of foods, nutrient values, yield factors… In these examples of recipes representatives of different types of dishes, we identified the main parameters affecting final result in recipe calculation: number and mass of ingredients in recipe, nutrient value and yield factor. Retention factors were intentionally not included in this list because EuroFIR made a high-quality synthesis in order to harmonize recipe calculation between food compilers [13]. These examples of recipes have shown us that choices of food compiler are an essential step and requires judgement and experience. Without a careful analysis of data, the risk is to obtain aberrant values, like in the case of vitamin C in the recipe of "Roasted vegetable chicken".

But this practical application has also demonstrated that this new method is a valuable tool in recipe calculation, allowing more precision. Indeed, when several choices are relevant (like 3 or 4 apples, or Bognar and USDA yield factor), mean and standard deviation of these two values allows to have a final result showing the range of possible values, not only one arbitrary value. Each time, the interval given around the medium value takes in account the values calculated with other choices (see fig. 1,2 and fig.5). Moreover, even when there is no choice to do, this method of calculation gives and uncertainty and variability due to standard deviation given in the original table. That is the case in all recipes, but examples of "Monkfish with mashed potatoes" and "Roasted broccoli rabe" are the more relevant about this point.

The knowledge of nutritional composition of food is essential for nutrition professionals and there are lots of factors of variability and uncertainty in this field [2]. Nevertheless, it is well known that recipe calculation is no more than an estimation [1][16]. It is in consequence essential to give a result accompanied by its variability and uncertainty, to give the consumer a range of values in which the true value has good chances to be. A far as we know, the method purposed in the present paper is the only one to do that.

This calculation method is based on differential calculus, according to the international method published by the GUM [6]. This is well adapted to recipe calculation because most of the time there is no more than two or three values per parameter in national databases. A collaboration with an experimental team would allow to have more values, and to improve our calculation model with Monte Carlo simulations. But that would imply to focus on a particular food. The advantage of the method presented in here is that its adaptable to any food database, according to the needs of users.

4. CONCLUSION

In conclusion, we have been developing a new method of recipe calculation, taking in account both variability between databases and the uncertainty of data, in one common calculus. This estimation of variability and uncertainty does not avoid doing a careful analysis of data before calculation, but would be a valued tool when accuracy is needed in recipe calculation and experimental data are not available.

Our first results have shown that it would be useful to make analysis of data before recipe calculation easier, by a systematic search of outliers among available nutritional data, using point plot and/or appropriate statistical test. This would allow to gain efficiency in recipe analysis and reduce bias due to food compiler influence.

REFERENCES

[1] W. M. Rand, J. a T. Pennington, P. Murphy, and J. C. Klensin, "Compiling Data for Food Composition Data Bases," *Nutrition*, pp. 1–65, 1991.

[2] J. A. T. Pennington, "Applications of food composition data: Data sources and considerations for use," *30th US Natl. Nutr. Databank Conf.*, vol. 21, no. Supplement 1, pp. S3–S12, 2008.

[3] Heli Reinivuo and K. Laitinen, "Proposal for the harmonisation of recipe calculation procedures: Composite Foods," 2007.

[4] D. A. T. Southgate and H. Greenfield, "Food composition data," p. 300, 2003.

[5] U. R. Charrondiere, D. Rittenschober, V. Nowak, C. Nicodemi, P. Bruggeling, and C. Petracchi, "FAO/INFOODS e-Learning Course on Food Composition Data.," *Food Chem.*, vol. 193, pp. 6–11, 2016.

[6] Joint Committee for Guides in Metrology (JCGM), "Evaluation of measurement data: Guide to the expression of uncertainty in measurement," no. September, p. 120, 2008.

[7] Allrecipes.com, "Roasted Vegetable Chicken Recipe." [Online]. Available: http://allrecipes.com/recipe/17676/roasted-vegetable-chicken/. [Accessed: 11-May-2016].

[8] Good Housekeeping - Easy Lamb Recipes, "Slow Cooker Indian Lamb." [Online]. Available:

http://www.goodhousekeeping.com/food-recipes/a13473/slow-cooker-indian-lamb-recipe-122652/. [Accessed: 11-May-2016].

[9] NYT Cooking, "Monkfish With Mashed Potatoes and Thyme Recipe." [Online]. Available: http://cooking.nytimes.com/recipes/11416-monkfish-with-mashed-potatoes-and-thyme. [Accessed: 11-May-2016].

[10] NYT Cooking, "Roasted Broccoli Rabe Recipe." [Online]. Available: http://cooking.nytimes.com/recipes/1016069-roasted-broccoli-rabe. [Accessed: 11-May-2016].

[11] NYT Cooking, "Free-Form Apple Or Pear Tart Recipe." [Online]. Available: http://cooking.nytimes.com/recipes/7573-free-form-apple-or-pear-tart. [Accessed: 11-May-2016].

[12] USDA, "National Nutrient Database for Standard Reference Release 28," *United States Department of Agriculture*, 2015. [Online]. Available: http://ndb.nal.usda.gov/ndb/foods. [Accessed: 13-Jan-2016].

[13] EuroFIR, "Report on collection of rules on use of recipe calculation procedures including the use of yield and retention factors for imputing nutri- ent values for composite foods," p. 184, 2008.

[14] A. Bognár, "Tables on weight yield of food and retention factors of food constituents for the calculation of nutrient composition of cooked foods (dishes)," 2002.

[15] B. A. Showell, J. R. Williams, M. Duvall, J. C. Howe, K. Y. Patterson, J. M. Roseland, and J. M. Holden, "USDA Table of Cooking Yields for Meat and Poultry Prepared by," pp. 1–30, 2012.

[16] EUROFIR, "HOW TO CALCULATE NUTRIENT CONTENT OF FOODS A GUIDELINE FOR FOOD BUSINESS OPERATORS," no. 1169, 2011.