Results of the Eurofoods trial on between-laboratory variation in the analysis of macronutrients in foods

Peter C. H. Hollman[1] and Martijn B. Katan[2]

[1] State Institute for Quality Control of Agricultural Products, Bornsesteeg 45, NL-6708 PD Wageningen, The Netherlands
[2] Department of Human Nutrition, Wageningen Agricultural University, De Dreijen 12, NL-6703 BC Wageningen, The Netherlands

Summary. In order to determine the influence of laboratory procedures on nutrient values in different food tables and data banks, an interlaboratory study was set up. Nineteen laboratories participated, and received well-homogenized samples of egg powder, full-fat milk powder, whole rye and wheat meal, biscuits and french beans to perform analyses of macronutrients by their own routine methods. For dry weight the results agreed very well; the results for ash agreed rather well. For protein the coefficient of variation between laboratories ($CV_{between}$) ranged from 2.8% to 6.4%. The $CV_{between}$ for total fat ranged from 5.4% to 54%. For available carbohydrates the $CV_{between}$ ranged from 9% to 27%. The $CV_{between}$ for total dietary fiber ranged from 23% to 84%. It is concluded that leading laboratories produce widely different values for macronutrients in common foods. Reference materials of certified nutrient concentration are needed.

Introduction

In order to improve the compatibility of nutrient data banks in Europe, Eurofoods has developed different activities [12]. The present trial was planned to determine the influence of difference in analytical and other procedures in laboratories that contribute to food tables, on the nutrient values in these tables.

Substantial information is available on the precision of specified analytical procedures, e.g. methods described by the Association of Official Analytical Chemists and the International Standards Organization. These data are collected by means of collaborative studies, in which all participating laboratories use the same accurately described method for the analysis of identical samples. However, it is a well-known fact that different laboratories actually use different methods to determine a certain nutrient in a certain food. Even if the same methodological principles are followed, subtile differences in procedure and in calibration materials could still cause large differences in outcome. Very few data are available about the influence of these differences. Therefore the Eurofoods subcommittee on laboratory analyses planned the present trial. Only the major macronutrients protein, total fat, available carbohydrates, total dietary fiber and ash were studied in products that can easily be homogenized and handled. One American and

Offprint requests to: P. C. H. Hollman

18 European laboratories that regularly contribute nutrient values to nutrient data banks participated in this study.

Methods

Materials

Six foods were selected for the trial:
1. Egg powder: commercially available spray dried whole egg powder.
2. Full-fat milk powder: commercially available spray dried full cream milk powder.
3. Whole rye: whole rye grains, donated by RIVRO-Institute (Wageningen, The Netherlands)
4. Whole wheat: whole wheat grains, donated by TNO/IGMB (Wageningen, The Netherlands).
5. Biscuits: Maria-biscuits (Koninklijke Verkade Fabrieken BV, Zaandam, The Netherlands).
6. French-beans: freeze-dried french beans (Summer Season, Coöp. Condensfabriek "Friesland" w.a., Leeuwarden, The Netherlands).

About 3 kg of each of these six foods were ground to pass a sieve of 0.5 mm mesh. The foods were carefully homogenized by quartering and divided into samples of ± 100 g, using a sample divider consisting of a rotary tube system, rotating at a frequency of $100 \, \text{min}^{-1}$. The samples were packed into airtight black plastic bottles with screwcaps. Prior to the distribution of the samples to the participants, sample homogeneity was tested as follows: Ten samples of each product were randomly chosen and each sample was analyzed for nitrogen. To determine the analytical precision, one sample of each product was also analyzed for nitrogen ten times. All analyses of one foodstuff were carried out in rapid succession by one analyst on one day. The results showed that there was no significant difference (*F*-test, 5%-level) between the coefficient of variation (CV) within the sample (analytical precision) and between the samples (CV < 0.25%). Samples could thus be regarded as homogeneous.

Statistical analysis

All results were calculated on dry matter as determined by each separate laboratory with a prescribed vacuum stove method. As all laboratories were asked to perform all analyses in duplicate with two technicians on different days, each providing one value, it was possible to calculate the variation within the laboratories. Statistical evaluation followed the principles of the International Standards Orga-

Table 1. Summary of the results of the dry weight determination

	Egg	Milk	Rye	Wheat	Biscuits	French beans
Number of labs.	15	15	15	14	15	14
	(g dry weight/100 g product as received)					
Mean	95.260	97.370	91.897	88.225	97.827	94.454
Range	94.6—96.5	96.9—98.6	91.3—93.0	86.7—89.2	97.4—98.5	93.3—95.3
CV_{within} (%)	0.18	0.12	0.13	0.18	0.08	0.13
$CV_{between}$ (%)	0.50	0.45	0.52	0.16	0.26	0.55

Table 2. Summary of the results for ash

	Egg	Milk	Rye	Wheat	Biscuits	French beans
Number of labs.	18	18	19	19	19	18
	(g/100 g dry weight)					
Mean	4.652	6.014	1.795	1.781	1.665	6.634
Range	4.3—5.8	5.7—6.7	1.6—2.0	1.6—2.1	1.5—1.9	5.9—7.8
CV_{within} (%)	1.8	0.7	3.0	2.1	4.4	1.6
$CV_{between}$ (%)	6.7	3.3	4.9	5.1	5.7	5.9

Table 3. Summary of the results of the protein determination

	Egg	Milk	Rye	Wheat	Biscuits	French beans
Number of labs.	17	18	19	19	19	18
	(g/100 g dry weight)					
Mean	52.983	28.100	10.175	12.662	7.840	15.092
Range	49.7—56.9	25.7—32.8	9.2—11.8	11.1—14.3	7.2—9.5	11.7—15.8
CV_{within} (%)	1.4	3.1	2.9	2.0	4.8	1.3
$CV_{between}$ (%)	2.8	5.2	6.4	6.4	6.2	6.2

nization norm ISO 5725-1981 [8] to calculate the standard deviations and coefficients of variation of within laboratory variation (s_{within}, CV_{within}) and between laboratories variation ($s_{between}$, $CV_{between}$). Contrary to ISO 5725, outliers were not rejected, because ISO 5725 only applies to interlaboratory tests with one method. Rejecting an outlying laboratory could imply rejecting a method that gives the "true" value. Moreover the aim of this interlaboratory trial was to investigate the influence of different laboratory procedures.

Results

Dry weight

The results of the dry weight determination by the prescribed vacuum stove method agreed very well, the coefficient of variation between laboratories ($CV_{between}$) ranged from 0.3% to 0.6% (Table 1).

Thus packing and storage conditions of the samples proved to be adequate to protect against changes in moisture content, and the prescribed method gives reproducible results.

Optional dry weight methods, mostly non-vacuum stove methods which some laboratories performed in addition to the prescribed method yielded results quite similar to the prescribed method.

Ash

Methods used show various pre-ashing procedures, ashing times (2—20 h) and temperatures (500—600°C).

The results for ash agreed rather well between laboratories, although outliers did occur (Table 2).
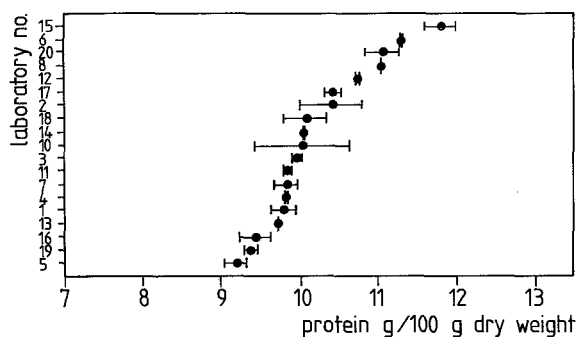
Protein

The results for protein show a $CV_{between}$ of 2.8% to 6.4% (Table 3). For a number of samples part of this variation is caused by the use of different Kjeldahl-factors. To eliminate the effect of these differences all results were recalculated using Kjeldahl-factors as recommended by FAO/WHO [4], which resulted in a slight decrease in variation between laboratories, especially with rye and wheat: 6.4% became 4.7% for rye and 6.4% became 5.2% for wheat.

The methods used differ in choice of catalyst, and procedures for digestion, distillation and determination of the ammonia formed. Most laboratories used $CuSO_4$ as a catalyst, three laboratories used selenium, three laboratories used mercury. Combinations of $CuSO_4$ and selenium (3

Table 4. Summary of the results of the total fat determination

	Egg	Milk	Rye	Wheat	Biscuits	French beans
Number of labs.	18	18	19	19	19	17
	(g/100 g dry weight)					
Mean	37.779	27.278	2.554	3.036	11.558	2.747
Range	29.4−44.2	24.5−30.0	1.6−4.5	1.8−5.8	9.9−15.4	1.2−5.8
CV_{within} (%)	2.0	2.0	24.5	30.6	2.7	25.7
$CV_{between}$ (%)	8.7	5.4	36.4	29.3	10.3	54.0

Fig. 1. Protein content of whole rye meal according to individual laboratories. |———| Duplicate values; ● mean value
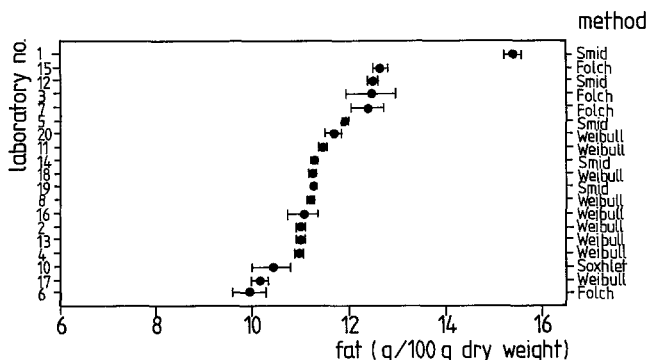
Fig. 2. Total fat content of biscuits according to individual laboratories. |———| Duplicate values; ● mean value

laboratories) and of $CuSO_4$ and TiO_2 (2 laboratories) were also used. The influence of the type of catalyst on the results for egg powder was investigated with the t-test. No significant differences ($P < 5\%$) were found so we did not examine the other products. Digestion was performed in block digestors, but classical Kjeldahl flasks were also used. Distillation of the ammonia was generally performed by steam distillation. Receiver solutions consisted of boric acid or sulfuric acid. Ammonia was mostly determined by titrimetric methods, sometimes using automated equipment. One laboratory used a colorimetric continuous flow method to determine the ammonia. Only one laboratory made a correction for nonprotein-nitrogen. This correction appears to have the greatest effect on values for milk powder and french beans. Precision data summarized in Table 3 lead to the following conclusions:

The discrepancies in protein values between laboratories are rather small, but are still higher than expected.

Variability for protein in cereals would be decreased if all laboratories used the Kjeldahl-factor of 5.83 recommended by FAO/WHO.

The difference between duplicate values is quite high in some laboratories (Fig. 1). The CV_{within} is higher than the values claimed by several participating laboratories (CV < 1%) for their own methods.

Total fat

The reproducibility of the fat determination was rather poor especially for products low in fat (Table 4). Thus, the fat content reported for whole wheat meal ranged from 1.8 to 5.8 g/100 g dry weight. For egg powder the range was 29 to 44 g/100 g.

Different methods were used. Acid hydrolysis followed by extraction with petroleum ether or diethylether was applied by most of the participants. Two different procedures were used: various methods according to Weibull-Stoldt [9] by which the fat after acid hydrolysis is separated by filtration, followed by a Soxhlet extraction, and various methods according to Schmid-Bondzynski-Ratzlaf [11] by which the fat after the hydrolysis is extracted by solvent partitioning. With all samples Schmid-methods gave on average higher results than Weibull-methods. These differences are significant ($P < 5\%$) for wheat, rye, biscuits and french beans. Milk powder was analyzed mostly by alkaline hydrolysis according to Röse-Gottlieb [9]. It appears that in milk powder Schmid-methods give higher results than Röse-Gottlieb-methods, which in turn give higher results than Weibull-methods. However these differences did not prove significant ($P < 5\%$). A number of laboratories used extraction techniques with different mixed polar and non-polar solvents such as chloroform/methanol and dichloromethane/methanol more or less similar to the Folch-method [5]. The performance of these methods with the different products is not quite consistent. Thus labs 3, 7 and 15 using the Folch-method, obtained high values for fat in rye and wheat meal, biscuits and french beans, but lab. 6, also using Folch reported lower values than average (Fig. 2). However in egg powder lab. 15 instead of lab. 6 found a lower-than-average fat content. The precision data for fat lead to the following conclusions:

The differences between laboratories in the fat content found in these foods are unacceptably high. Only part of this variability is due to differences in methods.

Within-laboratory variations were relatively large for products low in fat.

Available carbohydrates

Available carbohydrates were defined as follows: the sum of free sugars (mono-, disaccharides and other oligosaccharides

Table 5. Summary of the results for available carbohydrates

	Milk	Rye	Wheat	Biscuits	French beans
Original results					
Number of labs.	16	16	16	16	15
	(g/100 g dry weight)				
Mean	34.724	69.645	69.294	75.239	42.398
Range	14.9−44.4	38.4−94.0	35.7−82.1	63.4−89.3	28.3−67.5
CV_{within} (%)	4.8	4.2	4.5	3.3	3.1
$CV_{between}$ (%)	19.4	20.4	17.5	9.3	27.3
Expressed as monosaccharides and "by difference" values eliminated					
Number of labs.	14	13	13	13	12
	(g/100 g dry weight)				
Mean	35.317	71.745	71.535	78.168	41.990
Range	14.9−44.4	42.6−94.0	39.7−82.1	70.4−89.3	31.5−67.5
CV_{within} (%)	4.8	4.4	4.8	3.4	3.3
$CV_{between}$ (%)	19.5	18.5	15.6	6.9	22.6

Table 6. Summary of the results for total dietary fiber

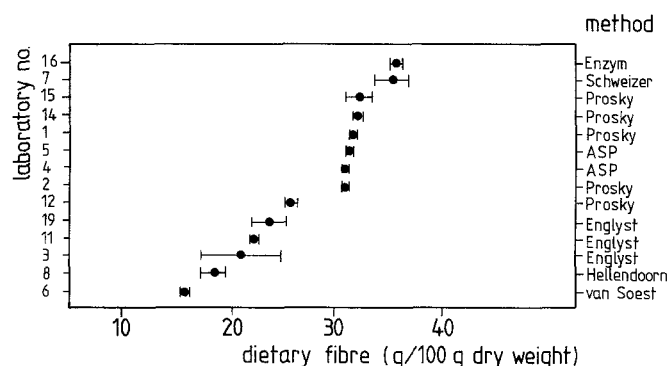|	Egg	Milk	Rye	Wheat	Biscuits	French beans
Number of labs.	4	7	14	14	14	14
	(g/100 g dry weight)					
Mean	0.361	0.278	15.427	13.109	3.116	27.610
Range	0−0.8	0−0.8	10.0−22.0	8.7−19.8	0.7−10.9	15.6−35.8
CV_{within} (%)	22.6	15.6	9.5	5.1	7.1	6.8
$CV_{between}$ (%)	115	129	24.6	26.4	84.0	22.8

Fig. 3. Total dietary fiber in french beans according to individual laboratories. |——| Duplicate values; ● mean value

up to approximately 10 monosaccharides units) and starch. Thirteen of the 19 laboratories performed carbohydrate analyses and three more calculated carbohydrates by difference. Seven laboratories expressed their carbohydrate results as monosaccharides, four as polymeric starch and two as "carbohydrates". The effect of these different modes of expression was investigated by recalculating all data to monosaccharides (Table 5). A factor of 1.11 was used to convert polysaccharides into monosaccharides. Few results have been reported for egg, so they are not shown in the table.

Methods used show many differences. Four laboratories isolated sugars and starch by separate extraction and deter-

mined in each extract sugars and starch with various methods. The sugars were extracted with water, 85% methanol or 80% ethanol. The other laboratories did not separate sugars and starch. Solubilization of starch was done by various techniques, e.g. DMSO, autoclaving or boiling with HCl or perchloric acid. Hydrolysis of the starch was performed mostly by enzymes such as amyloglucosidase or α-amylase. Sugars were determined by enzymatic, colorimetric, gas-chromatographic or reductiometric techniques. All these procedures were combined in different combinations. Three laboratories did not use an analytical method to determine the content of carbohydrates, but calculated this value by difference. This produced values close to the mean of the other laboratories, except for one laboratory whose results were generally higher. These high results can be expected because this laboratory determined crude fiber instead of dietary fiber. As a result, certain fiber component were counted as carbohydrate.

Expressing the results as monosaccharides (equal units) and omitting results calculated by difference, improves the precision slightly (Table 5).
It is concluded that:

The reproducibility of the available carbohydrate determination between laboratories was very poor.

Only a small part of the variability is due to different modes of expression, e.g. starch as monosaccharides versus starch as polymer weight.

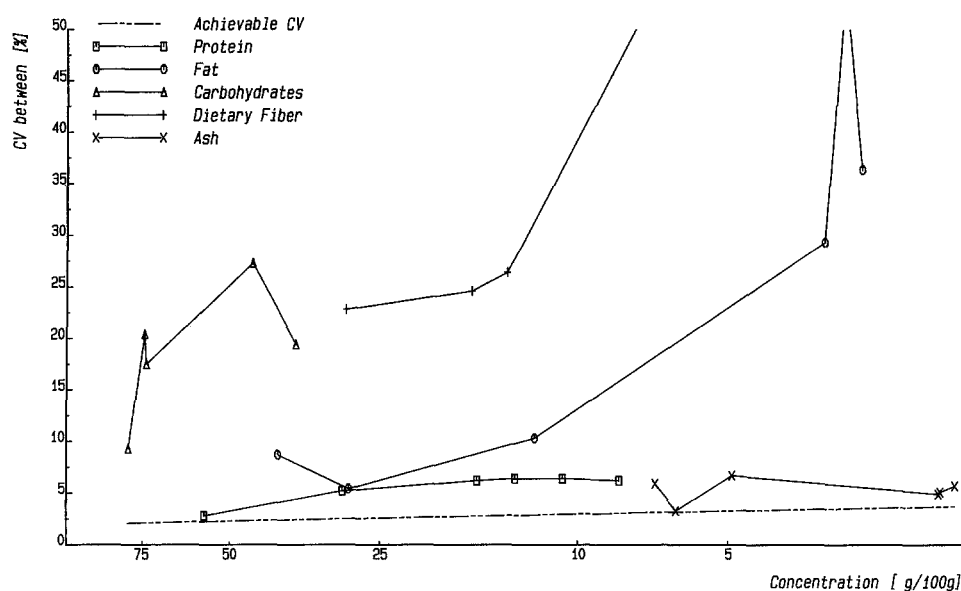Differences in methodology probably explain some but not all of the variability.

Fig. 4
Real $CV_{between}$ as a function of nutrient concentration in this trial, and theoretical CV achievable, when all laboratories use the same rigidly defined standardized methods [6]

Calculation of carbohydrates by difference causes no major bias, except when crude instead of total dietary fiber is used.

Total dietary fiber

Fourteen laboratories reported values for total dietary fiber (Table 6). Two more had determined crude fiber; these values were not used in the statistical analysis, because crude fiber is a small and variable part of total dietary fiber. The candidate AOAC-method described by Prosky et al. [10], used by 5 laboratories, and the related method described by Asp et al. [1], used by two laboratories (Fig. 3), resulted in values that agreed well. Labs 3, 11 and 19 used the Englyst-method [3] or a modification of it. Although the Youden rank test did not yield low outliers, labs 3, 6, 8, 11 and 19 did tend to report lower dietary fiber values than the trial mean. Prosky [10] also found that the Englyst-method gave lower values than the candidate AOAC-method, and pointed out that dietary fiber as determined by the Englyst-method does not include lignin. The low values of lab. 6 can be explained, because this laboratory used the neutral detergent fiber method, which determines only the water-insoluble fiber components.

In an interlaboratory study [10] recently organized to test the candidate AOAC-method, a CV for whole wheat of 11% was found as opposed to 27% in our trial where a variety of methods was used.

These data lead to the following conclusion:

There was a large variability in dietary fiber values as reported by different laboratories. This was probably due to well-known differences between methods.

Discussion

Results of this trial

The aim of this trial was to determine whether laboratory procedures could be a serious cause of discrepancies between different nutrient data banks in Europe. The trial has shown that this may indeed be the case. Leading food research laboratories in various countries produce widely different values for the concentration of fat, carbohydrates and fiber, and to a lesser extent also of protein, in everyday foods.

It should be noted that several sources of error that occur commonly in routine analyses of foods, had already been reduced or eliminated beforehand in this trial. Thus the foods were supplied as stable, well-ground powders of uniform particle size, easy to store, handle and sample. Also the samples had been carefully packaged and clearly marked and identified. Finally, the samples may have been analyzed with more than routine care. Because of all this, values produced in daily routine analyses of unknown samples will probably show an even larger variation between and within laboratories than the values reported here.

Causes of variability

As for the causes of those discrepancies, differences in methods probably play an important role. Thus Elkins [2] reported a much lower interlaboratory variability for protein and fat in the cooperative study of the Committee of Canning Industry Chemists. A main difference with the present study was that the participants in the Canning Industry study all used the same methods, as defined and described by the Association of Official Analytical Chemists, AOAC [7].

Horwitz [6] analyzed more than 150 collaborative studies (participants using exactly the same methods) organized by the AOAC, and was able to derive an empirical equation that relates the between-laboratory variation to the concentration (C) of the analyte, independent of the nature of the analyte or the analytical technique: $CV_{between} = 2^{(1-0.5 \log C)}$. It represents the reproducibility that can be obtained when all laboratories use the same rigidly defined standardized methods. As can be seen in Fig. 4 this calculated achievable $CV_{between}$ showed much lower values than the real $CV_{between}$ in the present trial, for fat, available carbohydrates, and total dietary fiber. So method effects were clearly visible in the present trial in the results for fat, available carbohydrates and fiber. However, differences in methodological principles are not the full explanation of the variability in results, as laboratories using similar methods

sometimes still reported widely diverging results. The cause for this is unknown.

Possible remedies

The trial has brought to light two types of variability. Firstly, within-laboratory variation was rather large for certain laboratories when analyzing certain nutrients and products. Such variability could be monitored and controlled by an adequate laboratory quality control programm.

Differences in level are responsible for most of the variability between laboratories observed in this trial. To minimize these differences better standardization of methods is needed. However, in food analysis standardization of methods is very complicated. In the mean time, differences in level can be detected by regular interlaboratory trials or by using external reference materials with a certified concentration of the nutrient of interest. As reference materials for macronutrients in foods are not available, this trial has shown that the production of such reference materials should have a high priority.

References

 1. Asp N, Johanssen C, Hallmer H, Siljeström H (1983) J Agric Food Chem 31:476−482
 2. Elkins E, Wolf W (ed.) (1985) Biological reference materials. John Wiley and Sons, New York, pp 357−363
 3. Englyst H, Wiggins H, Cummings J (1982) Analyst 107:307−318
 4. FAO/WHO (1973) Energy and protein requirements. Report of a Joint FAO/WHO Ad hoc Expert Committee. FAO Nutrient Meeting Report Series, No. 52, Rome
 5. Folch J, Lees M, Stanley G (1957) J Biol Chem 226:497
 6. Horwitz W (1982) Anal Chem 54(1):67A
 7. Horwitz W (1984) Official methods of analysis of the Association of Official Analytical Chemists, 14th edn. AOAC, Washington
 8. ISO 5725-1981. Precision of test methods − Determination of repeatability and reproducibility by inter-laboratory tests
 9. Osborne D, Voogt P (1978) The analysis of nutrients in food. Academic Press, London
10. Prosky L, Asp N, Furda I, De Vries J, Schweizer T, Harland B (1984) J Assoc Off Anal Chem 67(6):1044
11. Schormüller J (1968) Handbuch der Lebensmittelchemie, III/1. Springer, Berlin Heidelberg New York, p 807
12. West C (ed.) (1985) Ann Nutr Metabol 29 (Suppl. 1): 1−72