PROCEDURES FOR IMPUTING VALUES

Linda P. Posati
U.S. Department of Agriculture

Procedures for calculating and imputing nutrient values are used as circumstances prescribe by USDA's Human Nutrition Information Service in preparing portions of nutrient data bases. Major nutrient data bases available from USDA include the following: The revised sections of Agriculture Handbook No. 8; releases of USDA's Nutrient Data Base for Standard Reference; and the newly prepared USDA Nutrient Data Base for Individual Food Intake Surveys (1985). These data bases and others available from USDA may contain three types of nutrient values--analytical values which are usually arithmetic or weighted means, calculated values, and imputed values.

The distinction between analytical and calculated values and between calculated and imputed values is not clear cut. A goal in preparing data bases is to derive values which reflect the year-round, nationwide contribution of nutrients from foods. To achieve this goal, analytical data are sometimes weighted with factors obtained from production or marketing statistics. Calculated values are not themselves averages of analytical data but are derived from analytical data using specific procedures. They are no less valid than the analytical data upon which they are based and the soundness of the calculation procedures. Imputed values are quite often calculated values but they carry a degree of uncertainty. Analytical data upon which imputed values are based may be inadequate or assumptions that are without a sound research basis may have been made during the calculation steps. Below are data for cholesterol for three foods as presented in revised AH-8.

Food	Cholesterol (mg/100g)		
	Mean	Standard error	Number of samples
Whole milk	14	0.17	113
Dried egg yolk	548	--	133
Fresh ham, rump, lean and fat, raw	66	1.54	--

The mean values shown for whole milk and dried egg yolk are the arithmetic means of 113 and 133 samples, respectively. The standard error is given for the milk value to provide an estimate of variability. The standard error for the egg value is not calculated because the means were based on data that included averages of more than one analysis and for at least one of those averages no estimate of variance could be obtained. Analytical data contained in AH-8 are usually presented in the manner shown here for the milk and egg items. Calculated values are usually presented as means only without inclusion of a standard error or number of samples. The presentation of data for the fresh ham, rump cut, represents a deviation from this approach. The cholesterol content of this item is calculated using analytical data for cholesterol in the rump separable lean, the rump separable fat, and physical composition data for proportions of lean and fat in the particular ham cut. The mean is calculated using four variables which are estimated independently of each other using different numbers of replicates. Because of this there is no intrinsic number associated with the standard error.

Six specific procedures for calculating nutrient values from analytical data are based on or employ the following: Analytical and physical composition data; the content of a nutrient in a nutrient fraction of similar forms of a food; weighting factors; regression analysis; retention and yield data; and recipes. A commonly used procedure involves the use of the second item listed. Examples of nutrient fractions of foods are the lipid or fat fraction; the total solids; moisture-free, fat-free solids; the milk solids-not-fat fraction of dairy products; and moisture-free, fat-free, ash-free solids. Similar forms of a food would include the separable lean tissue from pork blade, rib, and top loin cuts. Similar forms of turkey are the flesh from the thigh, drumstick, and back, all of which are dark meat cuts. The pork cuts, on a total edible basis, would not be considered to be similar because their proportions of lean and fat differ, and the turkey cuts, if with skin, would not be similar because proportions of skin on the cuts differ.

A typical calculation based on analytical data for nutrients and physical composition data for cholesterol in raw chicken thigh with skin is as follows:

Tissue	Weight % of cut (EP)	Cholesterol (mg/100g)
Thigh meat	73.0	83
Skin	16.1	109
Separable fat	10.9	58

Cholesterol in cut =

(83 x .730) + (109 x .161) + (58 x .109) = 84 mg/100g

The cut consists of three tissues-thigh meat, skin, and separable fat. The mean percentages of the tissues are listed as weight percent of cut, edible portion. In the right hand column are listed means of the analytical data for cholesterol for each tissue. A calculated value for the chicken thigh with skin is obtained by multiplying the cholesterol content of each component tissue in the cut by the weight percent of each tissue and summing the results.

Below is a variation of calculations based on physical composition in which the analytical data are those for a nutrient in a nutrient fraction of the tissue. The value to be calculated is the grams of oleic acid per 100 grams of flesh and skin of a raw, whole chicken without neck and giblets.

Tissue	C18:1 (g/100g fat)	Fat %	Weight % of cut (EP)
Light meat	20.50 (27)*	1.65 (26)	31.95 (256)
Dark meat	25.84 (24)	4.31 (25)	39.49 (256)
Skin	34.22 (33)	32.34 (36)	17.09 (256)
Separable fat	37.22 (70)	67.95 (40)	11.45 (256)

% fat in flesh and skin:
By analysis = 15.06 (82)
Calculated = 15.54

C18:1 = [(20.50 x .0165 x .3195) + (25.84 x .0431 x .3949) + (34.22 x .3234 x .1709) + (37.22 x .6795 x .1145)] x (15.06/15.54) = 5.17 g/100g Food

*Numbers in ( ) = number of samples

Most of the fatty acid data compiled for poultry were obtained from analyses conducted on separated tissues which consisted of the breast and thigh muscle, skin, and separable fat. The four tissue components of the whole chicken are the light meat (breast and wing meat), dark meat (drumstick, thigh, and back meat), and the skin and fat. The second column contains normalized means for oleic acid on a gram-per-100 grams fat basis. Data for breast meat are used to represent light meat and data for thigh meat to represent dark meat. The oleic acid data were obtained as weight percent of methyl esters and converted to a gram-per-100 grams of fat basis by applying factors to account for the non fatty acid constituents of the lipid. The third column contains the means of analytical data for fat content of the four tissues. Percent fat for light meat is that for breast and wing meat in natural proportions and percent fat for dark meat is that for drumstick, thigh, and back meat in natural proportions. Samples analyzed for oleic acid and fat are not the same. The weight percent of each tissue in the cut represents a separate data set. The percent fat in raw chicken flesh and skin is 15.06%. A calculated fat value for flesh and skin obtained using figures in the two right hand columns would be 15.54%. The calculated oleic acid value for this particular cut is the sum of the products of oleic acid, fat, and weight percent of the four tissues adjusted by the ratio of analyzed to calculated fat content. Thirteen independent sets of analytical values are used to derive the value of 5.17 grams of oleic acid in the food.

A simple example of a calculation based on similar forms of food is for calcium content of 1% fat lowfat milks.

Known nutrient content: 13.88 mg calcium per gram milk-solids-not-fat (MSNF) of whole milk

Lowfat milk	MSNF (%)	Calcium (mg/100g)
1% fat	8.86	123 (13.88 x 8.86)
1% fat with added MSNF	9.20	128 (13.88 x 9.20)
1% fat, protein fortified	10.23	142 (13.88 x 10.23)

The mean calcium content per gram of milk solids-not-fat of whole milk is 13.88 mg based on a large body of analytical data for pooled milk. The nutrient contribution of different forms of milk depends primarily upon the level of milkfat and milk solids-not-fat in the products. Lowfat milks may be of three types-plain, with added MSNF, and protein fortified. The latter contains over 10% MSNF in the final product. A large amount of information on the percent of milk solids-not-fat in commercial lots of the three types of 1% fat lowfat milks was available. Very little analytical data were available for calcium in these milks as described. Rather than use the means of few data to represent the calcium levels in these foods, levels were calculated as shown using means of data for calcium in whole milk and for the milk solids-not-fat content of the three products.

HNIS sponsored research to obtain data on turkey for use in revising food composition tables for this commodity. Data were collected on proximate components, minerals, and four vitamins. Analyses were conducted on three turkey classes-fryer-roasters, young hens, and young toms. Vitamin B_6 was not among the nutrients examined. Levels in raw dark meat for each turkey class were calculated as shown in the following example:

Known nutrient content: B_6 in raw thigh meat of 18 young hens
0.351 mg B_6/100g wet tissue
20.22 g moisture-free, fat-free solids (MFFS)
0.0173 mg B_6/g MFFS

Turkey class	Number	MFFS (%)	B_6 (mg/100g)
Fryer-roaster	6	21.34	0.37 (0.0173 x 21.34)
Young hen	16	21.02	0.36 (0.0173 x 21.02)
Young tom	20	20.97	0.36 (0.0173 x 20.97)

Information from published literature was located for B_6 content in raw thigh meat of 18 young hens. The mean of these analytical values was 0.351 mg B_6/100g wet tissue. These particular samples contained an average of 20.22 grams of moisture-free, fat-free solids. The mean B_6 level in the analyzed samples was 0.0173 mg/g MFFS. The percentages of MFFS in the dark meat (flesh from back, drumstick, and thigh) are listed. The number of samples analyzed represents composites from 60 fryer-roasters, 60 hens, and 80 toms. The B_6 content for the dark meat of these three turkey classes was calculated by multiplying the known B_6 content of 0.0173 mg/g MFFS by the percentage of MFFS in the respective tissues. Differences in vitamin content between light and dark meat of poultry have been noted. However, effect of age or sex on vitamin levels is not well defined. If B_6 levels should be influenced by these factors, calculated values shown for fryer-roasters and young toms would be estimates.

The basic nutrient and physical composition data that HNIS had available for turkey were for the three classes previously mentioned. Nutrient values for a fourth classification of turkey, designated as "all classes" turkey, were calculated for use when the specific class was unknown and to provide a single set of nutrient values that would most appropriately reflect the nutrient contribution of turkey to our food supply. Below is an example of a procedure for deriving weighting factors for use in calculating nutrient values.

Turkey class	% of lb r-t-c (AP)		% EP in bird/100		Lb EP	Lb EP as %
Fryer-roaster	5	x	.7536	=	3.77	4.76
Young hen	45	x	.7836	=	35.26	44.53
Young tom	50	x	.8030	=	40.15	50.71
Sum					79.18	

Turkey class	Flesh and skin % of EP of bird		% edible contribution per class/100		Edible contribution by class to flesh and skin	Factors as %
Fryer-roaster	90.73	x	.0476	=	4.32	4.68
Young hen	92.08	x	.4453	=	41.00	44.40
Young tom	92.75	x	.5071	=	47.03	50.92
Sum					92.35	

The weighting factors used to derive the values for "all classes" turkey are based on the total pounds of ready-to-cook turkey of each of these three classes held in frozen storage over a certain time period. The proportions by weight contributed by the three classes are 5% fryer-roasters, 45% young hens, and 50% young toms. These values are for the as-purchased form of turkey including bone. The edible portion of each turkey class differs. Fryer-roasters contain more bone than do hens and toms. The pounds of edible portion contributed by each class expressed as a percent are shown in the right column with the top set of lines.

In order to use these factors to calculate specific cuts, the percentage of edible portion of each cut is taken into account. The first column on the lower set of lines contains the percent of edible portion contributed by the flesh and skin of each class. The difference between these values and 100 represents the edible portion of the neck and the giblets. To obtain weighting factors to calculate flesh and skin for an "all classes" turkey, these values are multiplied by the percent edible contribution per class as a decimal. The products expressed as a percent are shown in the lower right hand column. The factors are applied to nutrient data for flesh and skin of each class to calculate the "all classes" cut.

A procedure recently used to calculate nutrient values using analytical data for beef involved use of regression analysis. Below is an example of a calculation of fat content for a weighted market average value for beef top round steak, separable lean, raw.

OBS	Marbling Score	Observed Fat (%)	Predicted Fat (%)
1	5	2.78	2.47
3	8	3.10	3.13
6	9	3.31	3.35
7	10	3.55	3.57
9	12	4.14	4.02
10	13	4.77	4.24
11	14	4.35	4.46
12	15	4.94	4.68
14	23	5.91	6.45
16	24	7.73	6.69
18	25	7.17	6.89
19	26	7.27	7.11

Weighted market average: Marbling score = 12
Fat content = 4.02 g/100

Available for use were data for top round steaks from 19 beef carcasses. Only selected values are shown. The marbling score was available for each animal. This score is the primary quality grade determinant. The quality grades assigned to animals with these marbling scores were as follows: Standard grade-scores 5-7; good-scores 8-11; choice-scores 11-19; and prime-scores 20-26. Regression analysis showed a high correlation of both fat and water with marbling score. The observed fat values for the top round steaks are shown in the third column in order of increasing marbling scores of the animals. The right hand column contains the predicted fat values for the marbling scores shown. Data collected by USDA over the period 1980-1985 on graded beef for mean marbling scores within quality grades reflected a weighted market average marbling score of 12, equivalent to a grade of low-choice. For an overall value representative of the market, a predicted fat value of 4.02 g/100g associated with the score of 12 is appropriate.

It has been our experience to find more analytical data available on raw forms of food than on cooked forms. Therefore, cooked values are frequently calculated from raw data using information on nutrient retentions and cooking yields. A sample calculation for thiamin in light meat without skin for a mature stewed chicken follows:

Analytical data:

Thiamin in raw meat = 0.132 mg/100g
Thiamin retention = 51%
Yield of cooked meat = 72%

Thiamin in cooked meat:

$$\frac{0.132 \times 0.51}{0.72} = 0.094 \text{ mg/100g}$$

Analytical data available are the thiamin content of the raw light meat of mature chicken; the percent retention of thiamin in light meat upon stewing based on cooking studies on broiler-fryer chickens; and the yield of cooked light meat from mature chicken upon stewing. The thiamin content of the cooked meat is calculated by multiplying the thiamin content of the raw meat, 0.132 mg/100g, by the retention fraction, and dividing the product by the cooking yield to express the value on a 100-gram basis.

HNIS compiles a considerable amount of data on cooking yields and the retention of nutrients for various cooking methods for use in these types of calculations. An example of the determination of cooking yield factors for chicken breast using one particular sample from a study on chicken that HNIS had sponsored follows.

	Left Breast	Right Breast		
	Raw Wt	Raw Wt	Roasted Wt	
	g	g	g	g x 194/192
Part	194	192	134	135
Meat	132		97	98
SKin	17		14	14
Fat	6		—	—
Bone	37		21	21

Part yield = (135 x 100)/194 = 70%
Meat yield = (98 x 100)/132 = 74%
meat + skin yield = [(98 x 14) x 100]/(132 + 17 + 6) = 72%

Anatomically matched cuts representing opposite sides of the same chicken were utilized. The breast from one chicken was split into left and right halves. The left raw breast was weighed at 194 grams and then separated into its component parts—meat, skin, separable fat, and bone and each were weighed. The right raw breast weighed 192 grams. It was then roasted, weighed after cooking at 134 grams, and finally separated into meat, skin, and bone. The right hand column contains the cooked weights adjusted by the ratio of the raw weights of the left and right breast to eliminate discrepancies due to cutting. The cooking yields are calculated using the data in the left and right hand columns. The cooked yield of the entire breast, that is the part with bone, equals 70%. The yield of cooked meat from raw meat is the weight of cooked meat, 98 grams, divided by the weight of raw meat, 132 grams, and equals 74%. The yield of the total edible portion of the breast is the sum of the weights of cooked meat and skin divided by the sum of the weights of raw meat, skin, and separable fat and equals 72%. The different tissues, including bone, lose different amounts of weight on cooking.

Cooking yield data are used in developing retention factors. The preferred method for calculating nutrient retentions is based on measuring the proportion of nutrient remaining in a cooked food in relation to the amount of that nutrient originally present in a given weight of the food before cooking. Shown is an example of changes in proximate components upon roasting of the chicken breast meat and skin sample used for the previous cooking yield calculations.

	Raw	Roasted	% Retention
Weight g	155	112	
Water %	70.01	62.89	64.9
Protein %	22.06	28.12	92.1
Fat %	7.28	7.45	73.9
Ash %	.99	.99	72.3

	Raw (g/100g)		Reten-tion	=	Cooked (g/72g)	Loss or gain (g)
Water	70.01	x	.649	=	45.44	-24.57
Protein	22.06	x	.921	=	20.32	- 1.74
Fat	7.28	x	.739	=	5.38	- 1.90
Ash	.99	x	.723	=	.72	- .27
Sum	100.34				71.86	28.48

The raw weight of 155 grams is the sum of the weights of raw meat, skin, and separable fat. The roasted weight of 112 grams is the sum of the weights of cooked meat and skin. The weight of 112 grams divided by the weight of 155 grams represents the cooking yield and is equivalent to 72%. Below the weight values are listed the analytical values for the raw and roasted tissue. The percent retention for water, for example, is the amount of water in the cooked breast with skin, 62.89%, multiplied by the 72% yield factor and divided by the percent water in the raw breast with skin, 70.01%. Nutrient retentions shown are about 65% for water, 92% for protein, 74% for fat, and 72% for ash. The proximate data for the raw sample are repeated in the first column on the lower set of lines. These values, when multiplied by the retention factors, yield the amounts of the four components in 72 grams of cooked product. The differences between these numbers and the values in the left hand column represent the changes in nutrients upon cooking. A 100-gram sample of raw chicken breast with skin would lose about 25 grams of water, 1.7 grams of protein into the drippings, 1.9 grams of fat, and 0.27 grams of ash to drippings. The solids lost on cooking, that is protein, fat, and ash, represent 14% of the total weight loss of 28 grams. Nutrient retentions calculated using data on a dry-weight basis are generally higher than those calculated using the actual yield or weights of raw and cooked foods because weight loss is presumed to be due to water loss alone. Such is not the case for meat products which lose fat to drippings.

The following examples of procedures for calculating nutrient values are based on recipes. An important aspect of any recipe calculation is the determination of ingredient proportions. Below is an example of development of a formulation based on a partial nutrient profile. The food item is a chocolate coated ice cream bar.

Analytical Data		Ingredient	NDB No.
Pro	3.4%	1 Vanilla ice cream	01061
Fat	22.3%	2 Coconut oil	04047
Carb	25.7%	3 Sugar	92300
Iron	0.2 mg	4 Cocoa, high-fat Dutch	77820
Vit A	308.0 IU		

Ingredient (%)		Calculated Values	
1	79.0	Pro	3.0%
2	13.5	Fat	22.3%
3	6.4	Carb	25.7%
4	1.1	Iron	0.2 mg
		Vit A	322.5 IU

Analytical data are shown for protein, fat, carbohydrate, iron, and vitamin A. Ingredients listed on the label for the product are those for ice cream and the chocolate coating. Coating ingredients in decreasing order are coconut oil, sugar, and high-fat Dutch cocoa. Each ingredient corresponds to an item contained in USDA's Nutrient Data Base for Standard Reference. The Nutrient Data Bank (NDB) numbers are shown for the item contained in the Standard

Reference data file. Using as input the analytical data shown for the five nutrients listed and the NDB numbers for the ingredients in decreasing order, the percentages of the ingredients are determined using a computer program which adjusts proportions of ingredients until calculated nutrient values converge as closely as possible with the analytical values. Use of the percentages of ingredients shown in the lower left hand column yields the calculated values shown on the right.

If the proportions of cooked ingredients in a mixed dish are known, nutrient values may be calculated for a mixture using data for cooked foods, thereby eliminating the need to apply retention and yield factors. Alternatively, recipe calculations may be based on raw ingredients. An example of a recipe calculation based on raw ingredients follows:

Egg dessert, baked	Food code:321-2010
Yield: 90.0	Fat change: 0.0
Moisture change: - 10.0	Fat ID:

NDB No.	Ingredient	Retention	Measure	Grams
92300	Sugar	0	3 c	600
01125	Egg yolk	101	16	272
01123	Whole egg	101	2	100
97010	Water	0	3/4 c	177

The format for the input information is that used with the HNIS recipe linking program system for automatic calculation which was described at this conference. The item for which nutrient values are to be calculated is a baked egg dessert. The food code for this item as contained in the Nutrient Data Base for Food Intake Surveys is recorded. The cooked yield of the product is 90%. Fat change is indicated as zero, fat being neither absorbed nor lost during preparation. Cooking losses for this item are due to evaporation. Moisture change is indicated as minus 10%. The NDB numbers for the four ingredients of the item are listed in the left hand column. A retention code is entered for two of the ingredients in this example. The code 101 is used to access retention factors for vitamins and minerals in egg cooked by a dry-heat method. These factors are contained in a separate data file. This particular recipe is the standard cook book type — 3 cups of sugar, 16 egg yolks, 2 whole eggs, and 3/4 cup of water. Both the measures and their corresponding gram weights are part of the entry information. The calculation steps are outlined below.

Ingredient	Weight (g)	Energy (Cal)	Water (g)	Fat (g)	B_{12} (mcg)
Sugar	600	2310	3.0	0	0
Egg yolk	272	1004	132.6	89.5	10.34 x .8
Whole egg	100	158	74.6	11.2	1.55 x .8
Water	177	0	177.0	0	0
Subtotals:	1149	3472	387.2	100.7	9.51
Mois/fat ch:	-115	0	-114.9	0	0
Yield:	1034	3472	272.3	100.7	9.51
Per 100 grams:	100	336	26.3	9.7	0.92

Nutrient values for energy, water, fat, and B_{12} on a 100-gram basis are accessed from the Standard Reference Data file. The values are converted to the amounts in the specific weight of each ingredient. Retention factors of 0.8 (80%) are applied to the B_{12} data. The values are summed as shown on the subtotal line. The next line shows the changes in weight and water content during preparation. Ten percent of the weight is deducted from both the sum of ingredient weights and the grams of water. The yield line contains amounts of nutrients in 1,034 grams of food. These values are converted to the per-100-gram basis. If these calculations were done with the recipe linking program system, the final values on the 100-gram basis would automatically be written to a data file.

A final example of a recipe calculation shows how changes in fat, in this case a loss of fat, may be taken into account. The item to be calculated is stewed turkey.

Turkey, stewed
Yield: 75.6
Moisture change: -21.1

Food Code: 242-0140
Fat change: -3.3
Fat ID: 04575

Ingredient	Weight (g)	Energy (Cal)	Water (g)	Fat (g)	Cholest (mg)
Turkey, flesh & skin, raw	100.0	160	70.4	8.0	68
Mois/fat change:	-24.4	-30	-21.1	-3.3	-3
Yield:	75.6	130	49.3	4.7	65
Per 100 grams:	100.0	172	65.2	6.2	86

The starting ingredient is the raw item. Fat loss is shown as 3.3%; moisture loss is 21.1%. The yield of product is 100 minus the sum of the fat and moisture loss and equals 75.6%. Fat identification, on the upper right, is the NDB number 04575 which is the code for turkey fat. Nutrient values for energy, water, fat, and cholesterol for 100 grams of raw turkey flesh and skin are listed on the first lower line. The second line shows adjustments made to the weight and nutrients to account for water and fat losses. The cooking loss of 24.4 grams is subtracted from the raw weight. Thirty calories from fat are deducted. Grams of water and fat lost are subtracted. The 3.3 grams of turkey fat contained 3 mg of cholesterol, which is also deducted. The amounts of the nutrients contained in 75.6 grams of cooked food are shown on the line for yield. These figures are converted to the 100-gram basis by division by 0.756.

The typical calculation procedures shown can also be used to impute nutrient values. Imputed values bear some degree of uncertainty. They may be calculated from label declarations for fortified foods and by using closely related but different foods or recipes with uncertain ingredient proportions. In some cases it may be necessary to use estimated retention factors, cooking yields, or physical composition. In order to impute values for certain nutrients it may be necessary to extrapolate data from a parent food to processed forms of the food when little information on the processing effects is available. Such factors affect the validity of the final results.