Journal of Food Composition and Analysis 22 (2009) 718–727

Contents lists available at ScienceDirect

Journal of Food Composition and Analysis

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

Original Article

Quality assurance with an informatics auditing process for Food Composition Tables

Chi-Ming Chu [a], Meei-Shyuan Lee [a,f,*], Yi-Hsien Hsu [a], Hsiao-Li Yu [a], Tsai-Yi Wu [a], Su-Chien Chang [a], Li-Ching Lyu [b], Fang-Ju Chou [c], Yun-Ping Shao [d], Mark L. Wahlqvist [a,e,f]

[a] *School of Public Health, National Defense Medical Center, 161 Minchuan East Road, Section 6, Taipei 114, Taiwan, ROC*
[b] *Department of Human Development and Family Studies, National Taiwan Normal University, 162, HePing East Road, Section 1, Taipei 106, Taiwan, ROC*
[c] *Department of Nutrition, Wan Haw Hospital, 6 Chunghaw Road, Section 2, Lane 606, Taipei 108, Taiwan, ROC*
[d] *Department of Food and Nutrition, Tri-Service General Hospital, National Defense Medical Center, 325 Chenggong Road, Section 2, Taipei 114, Taiwan, ROC*
[e] *Institute of Population Health Sciences, National Health Research Institutes, 35 Keyan Road, Zhunan Town, Miaoli County 350, Taiwan, ROC*
[f] *Asia Pacific Health and Nutrition Centre, Monash Asia Institute, 8th Floor, South Wing Menzies Building (11), Monash University, Melbourne 3800, Australia*

ARTICLE INFO

Article history:
Received 7 August 2008
Received in revised form 22 January 2009
Accepted 2 March 2009

Keywords:
Quality assurance
Auditing process
Food composition database
Coefficient of variation
Food subgroups
Compositional similarity
Food variety
Taiwan
Food composition

ABSTRACT

A 6-step auditing process was developed to detect unlikely nutrient values in a Nutrient Composition Data Bank for Foods (NCDBF) in Taiwan. Preference was given to finding errors in the database, rather than to determining significant differences in the biological characteristics of the individual nutrients. There were 239 compositionally similar subgroups categorized within the NCDBF. The coefficient of variation (*CV*) of nutrient values for each subgroup provided the first-order sorting instrument. Nutrient *CV*s were ranked in rows for food subgroup (x) and in columns for nutrient type (y) and their product (x,y) in descending order. When the rank was in the top 2 or the product was ≤ 20, the Excel "cell" was regarded as a "*hit*". The "*hit rate*" (2.6%, 777 *hits*/29,424 pieces of information) of the computerized analysis was verified through an expert panel review to provide a "*satisfied hit rate* (*SHR*)" (agreed errors/total food group *hits*). The mean *SHR* was 14.9% (range: 1.4%–37.6%) for the various food groups. The computerized process performed with a 38-fold increase in likelihood of error detection compared with what manual assessment alone would have produced. This low-cost approach could be applied in various jurisdictions or with other digitized food composition tables.

1. Introduction

Food Composition Tables (FCT) are used to translate data on the nutrients in food for use in labelling, nutrition education, dietary surveys, epidemiological studies, and dietary intervention trials (Bowers, 2002; Buzzard et al., 1990; Deharveng et al., 1999; Dwyer, 1994; Fidanza and Perriello, 2002; Guilland et al., 1993; Leclercq et al., 2001; Matsuda-Inoguchi et al., 2004; Sasaki et al., 1999; Shimbo et al., 1996; Vaask et al., 2004; Zhang et al., 1999). FCT comprise the most frequently used and representative foods in a region, and contain information on energy and generally dozens of nutrients in numerous foods. The quality of food composition information is dependent on the sampling of foods, the analytical procedures used and how data are presented. Errors can have a profound effect on research findings and food judgements, especially where the information is designed to connect with public health and nutrition education initiatives. Of particular concern is the accuracy of nutrition information in food labelling, which depends on reliable food compositional data (Fabiansson, 2006).

The nature and derivation of nutrient values in food items is often misunderstood, even among specialists, with consequent error in reporting nutrient intakes. For instance, in the Taiwanese FCT, mung bean sprouts have the highest vitamin C per 100 g dry weight among vegetables; yet soybean and mung bean sprouts contain 13 mg and 184 mg vitamin C per 100 g dry weight, respectively (Taiwan Department of Health, 2002c). The difference is 171 mg or 14-fold. Such discordant values between two apparently similar foods create difficulties in the accurate estimation of individual or community nutrient intakes.

Although such discrepancies are clearly recognized, it is not feasible to inspect the quality of FCT data using traditional laboratory methods alone; they are too expensive. This is especially true when assays must be repeated or when this applies to more than 20,000 data points (i.e. assuming 20 nutrients per

* Corresponding author at: School of Public Health, National Defense Medical Center, 161 Minchuan East Road, Section 6, Taipei 114, Taiwan, ROC.
Tel.: +886 2 87910704; fax: +886 2 87910704.
E-mail address: mmsl@ndmctsgh.edu.tw (M.-S. Lee).

0889-1575/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jfca.2009.03.005

1000 food items). A simple and low-cost auditing process for identifying unlikely nutrient values in foods in FCT is needed so as to enable experts to review the values and suggest either more selective revision, or recommend that another laboratory analysis be performed. Such an auditing process would represent significant savings in time, costs and human resources; it would require computer power which would remove limitation by way of the size of datasets, a situation which might be encountered in any future FCT.

Computers and information technology have become tools to improve efficiency and quality in every profession. Nutritional values can be evaluated through an automated process (Alexander et al., 1993; Dartois et al., 1989; Probst et al., 2004; Samsonov et al., 1986) since informatics applies to the work of nutritionists in public health and individual health (Mccullough et al., 1999; Phillips et al., 2006; Rodriguez et al., 1995; Sharp and Ahmed, 1983; Van Wave and Decker, 2003). A robust approach can efficiently scrutinize nutrient values of food items in several datasets (NATA, 1996; ISO/IEC, 1997a; ISO/IEC, 1997b; Puwastien, 2002). Other approaches include the recognition of extreme medians and normalized inter-quartile ranges by use of the *z*-score (Institute for Interlaboratory Studies, 2006), where data with *z*-scores outside of the range of ±3 were not acceptable. The International Network of Food Data System (INFOODS) (Rand, 1985; Rand, 1987; Rand et al., 1991; Rand and Young, 1984) and the Association of South East Asian Nations Network of Food Data Systems (ASEANFOODS) (Puwastien, 2002) use this approach to approve pooled datasets of FCT, as have the Greenfield and Southgate (Southgate and Greenfield, 1992) guidelines.

Nutritionally similar foods should have similar nutrient values. However, foods may be grouped at different levels, such as fruits then citrus, vegetables then dark green vegetables. For educational purposes, foods are not introduced to the public singly, but usually categorized by their biological similarity, e.g. citrus fruits in the fruits group and dark green vegetable in the vegetables group are rich in vitamin C and carotene, respectively. Each of the food subgroups, citrus fruits or dark green vegetables, would be expected to be more nutritionally similar than their parent groups, fruits and vegetables, in accord with their known food biology whose features inform dietary guidelines and food guides. Although the conventional manual approach to the evaluation of FCT quality might discover extreme nutrient variation in a food group, it may neglect unlikely nutrient values for a subgroup of similar food items within it. Neglected values are sources of potential error in FCT, and cannot be identified by the *z*-score for a food group as a whole.

There is a demand for reliable food composition information that can be validated through computerized dietary system tools (Farran-Codina et al., 1994; Shai et al., 2003). We propose an informatics process for data quality audit to identify unlikely nutrient values. The underlying premise is that within a food subgroup, characterized by biological similarity of its food items as basic commodities or by way of recipe or food process involved in preparation, an outlier in nutrient composition might be an error which merits revision.

2. Materials and methods

2.1. Nutrient composition data bank for foods in the Taiwan area

The Nutrient Composition Data Bank for Foods (NCDBF) in the Taiwan Area consists of 24 nutrients (including energy and water) in 1226 foods of 18 groups like CEREALS (119 items), POTATOES & STARCHES (17 items), and NUTS & SEEDS (39 items). The number of food items for each food group is shown in the second column of Table 1 (Taiwan Department of Health, 2002a; Taiwan Department of Health, 2002b).

Table 1
Audit of Taiwan Food Composition Tables by statistical "*hits*" and an Expert Panel to identify unlikely nutrient values in 18 food groups.

Food Group	Computerized process						Panel process		SHR(%)[a]	
	All food items		Unprocessed food items		Sum of hits	HR/ PER (%)[b]	Panel found missing or was satisfied		Without MH	With MH
	Items[c]	Subgroups[d]	Items[e]	Subgroups[f]			MH[g]	SH[h]		
CEREALS	119	23	16	6	49	1.72	0	1	2.04	2.04
POTATOES & STARCHES	17	5	11	4	–		–	–		–
NUTS & SEEDS	39	11	25	10	57	6.09	0	11	19.3	19.3
FRUITS	101	22	70	21	72	2.97	0	23	31.9	31.9
VEGETABLES	138	29	67	28	82	2.48	6	27	32.9	30.7
ALGAE	5	1	5	0	–		–	–		–
MUSHROOMS	17	7	11	6	–		–	–		–
PULSES	69	13	45	12	68	4.11	0	5	7.35	7.35
MEATS	101	25	88	24	89	3.67	4	35	39.3	37.6
FISH & SEAFOODS	103	18	66	17	–		–	–		–
EGGS	31	7	20	6	47	6.32	5	4	8.51	7.69
MILKS	66	13	41	12	66	4.17	1	6	9.09	8.96
FATS & OILS	42	13	40	12	–		–	–		–
SUGARS & SWEETENERS	10	2	2	1	–		–	–		–
BEVERAGES	86	7	15	2	–		–	–		–
SEASONINGS & SPICES	82	12	38	11	70	3.56	1	1	1.43	1.41
CONFECTIONERIES	64	9	0	0	69	4.49	1	1	1.45	1.43
PROCESSED FOODS	136	22	0	0	90	2.76	0	2	2.22	2.22
Total	1226	239	560	172	759	2.58	18	116	15.3	14.9

[a] *SHR*: *satisfied hit rate* = (satisfied *hits*/sum of hits) × 100%.
[b] *HR* (or *Probable error rate, PER*): *hit rate* = (sum of *hits*/total piece of information) × 100%, where total pieces of information are the number of foods multiplied by the number of nutrients, which is 24 in the NCDBF.
[c] Number of foods in each food group in all foods.
[d] Number of subgroups plus one unsorted subgroup in all foods.
[e] Number of items of unprocessed foods (non-treated or non-prepared).
[f] Number of subgroups with items of unprocessed foods.
[g] *MH*: missing hit, unlikely values were identified by the Panel, but not by the computerized process.
[h] *SH*: *satisfied hits*, system-defined *hits* which satisfied Expert Panel as legitimate errors worthy of further evaluation.

2.2. Terminologies, abbreviations and algorithms or formulas

(1) Coefficient of variation, *CV*: Statistical measure of dispersion in a data series around the mean. It is defined as the ratio of the standard deviation to the mean. In this process, the population standard deviation (σ) is calculated because there is no sampling. ($CV = \sigma$/mean).
(2) *CV RANK*: Order of descending *CV*s by nutrients and food subgroups, e.g. it is 1 for the largest *CV*. (x or y for subgroup or nutrient, respectively).
(3) *CV RANK* Product, CRP: Product of *CV RANK* (x,y). (CRP = $x \times y$).
(4) Outlier: Individual value which is far away from mean, i.e. mean ± 1.5 SD (standard deviation) and mean ± 3.0 SD.
(5) Hit: Computerized, process-defined unlikely value.
(6) Hit rate, *HR*: *Hits* per total pieces of information in each food group or in each dataset which is a ''probable error rate'' (*PER*). *HR* = (sum of *hits*/total piece of information) × 100%, where total piece(s) of information is the number of foods multiplied by the number of nutrients, namely 24 in the NCDBF.
(7) Missing hits, *MH*: unlikely values were identified by the Panel, but not by the computerized process.
(8) Satisfied hit, *SH*: system-defined *hit* with which the Panel agrees
(9) Satisfied hit rate, *SHR*: the proportion of *Satisfied hits* to total *hits* which is calculated as (*Satisfied hits*/Sum of *hits*) × 100%

2.3. The audit process

We developed an audit process to identify unlikely nutrient values in food items using the following steps (Fig. 1):

(1) *Conversion of nutrient content to a dry weight (DW) basis, i.e. mass of nutrient per 100 g dry food*: to adjust for the impact of water on nutrient content and deal with a largely irrelevant source of variability. Thus [original amount of a nutrient per 100 g fresh wt/(100-amount of water in 100 g fresh wt) × 100] was used for nutrient content per 100 g dry food.
(2) *Assignment of food items to a subgroup*: to increase the degree of similarity, particularly for processed food, firstly whether it was ''processed'' or ''unprocessed'' and, then the method of preparation such as pickled, dried or canned to constitute various subgroups. (See Appendix A for examples.)
(3) *Calculation and ranking of CVs:* to calculate the *CV*s of nutrient values of food items in one subgroup and to rank *CV*s, in descending order, by nutrients and food subgroup, respectively. (Hereafter referred to as ''*CV RANK*'')
(4) *Identification of unlikely values*: to find outliers by a filter which uses the subgroup nutrient *CV*s, being the top 2 ranks in each subgroup for either the subgroup (x) or its nutrient (y); or the product of the *CV RANK*s x and y as a vector (x,y), defined as CRP, with a value less than 20. Each separate vector is referred to as a ''cell'', which provides the basis for a ''*hit*'' (refer to Table 2 and Section 3.4 for the detail of the filter).
(5) *Judgment by a food and nutrition expert panel*: to verify the findings and to ascertain whether the filtering method missed unlikely values. Experts who had formal training in the food and nutrition sciences and had achieved professional distinction were invited.
(6) *Recommendation for further work*: to accept, to re-analyse or to revise the nutrient of a given food as required.

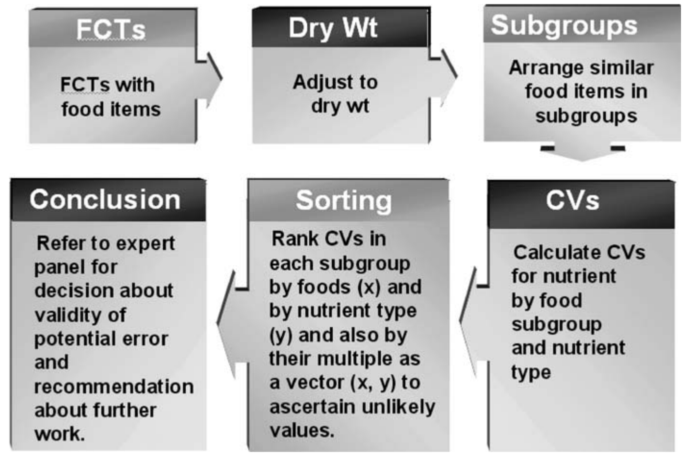


Fig. 1. Six-step informatics audit process.

Table 2
Empirical leverage of food (x) or nutrient (y) *CV RANKs* and their product (x,y) for the filter of the VEGETABLES group.

CV RANKs (x or y)[a]		*CV RANK* Product (x,y)[b]								
		0	5	10	15	20	25	30	35	40
0	*SH*[c]	0	8	18	23	26	27	27	27	27
	SHR (%)[d]	–	32	36	34.3	32.5	29.7	27	24.8	22.1
1	*SH*	14	14	19	23	26	27	27	27	27
	SHR (%)	37.8	35.0	36.5	33.8	32.1	29.7	27.0	24.8	22.1
2	*SH*	21	21	21	24	27[e]	27	27	27	27
	SHR (%)	32.8	32.8	32.8	30.8	30.7	28.4	26.5	24.8	22.1
3	*SH*	24	24	24	24	27	27	27	27	27
	SHR (%)	25.3	25.3	25.3	25.3	26.5	25.7	25.0	23.9	21.6
4	*SH*	27	27	27	27	27	27	27	27	27
	SHR (%)	22.3	22.3	22.3	22.3	22.3	22.1	22.0	21.8	20.6
5	*SH*	28	28	28	28	28	28	28	28	28
	SHR (%)	19.6	19.6	19.6	19.6	19.6	19.6	19.6	19.6	19.2

[a] The descending rank of coefficient of variation (*CV*) of either the food (x) or its nutrient (y).
[b] The product of the rank of *CVs* of (x) and (y).
[c] *SH*: satisfied hits, system-defined hits which satisfied Expert Panel as legitimate errors worthy of further evaluation.
[d] *SHR*: satisfied hit rate = (satisfied hits/sum of hits) × 100%.
[e] The rank 2/product 20 cell of 27/30.7 represents the filter for preferred screening threshold for *hits*.

3. Results

We demonstrate the nature and detail of this six-step audit process by using the NCDBF.

3.1. Conversion of nutrient content to a dry weight basis

Nutrients in foods were compiled into individual Excel worksheet by groups. The example of the VEGETABLES group is shown in Fig. 2. The columns present the codes of food in "A", the subgroup of food items classified as similar in "B", whether processed or not in "C", the name of food in "D", the energy in "E", the water in "F", and the contents of 22 nutrients from "G" (protein) to "AA" (zinc). The nutrient contents as DW were derived to take water content into account.

3.2. Assignment of food items to a subgroup

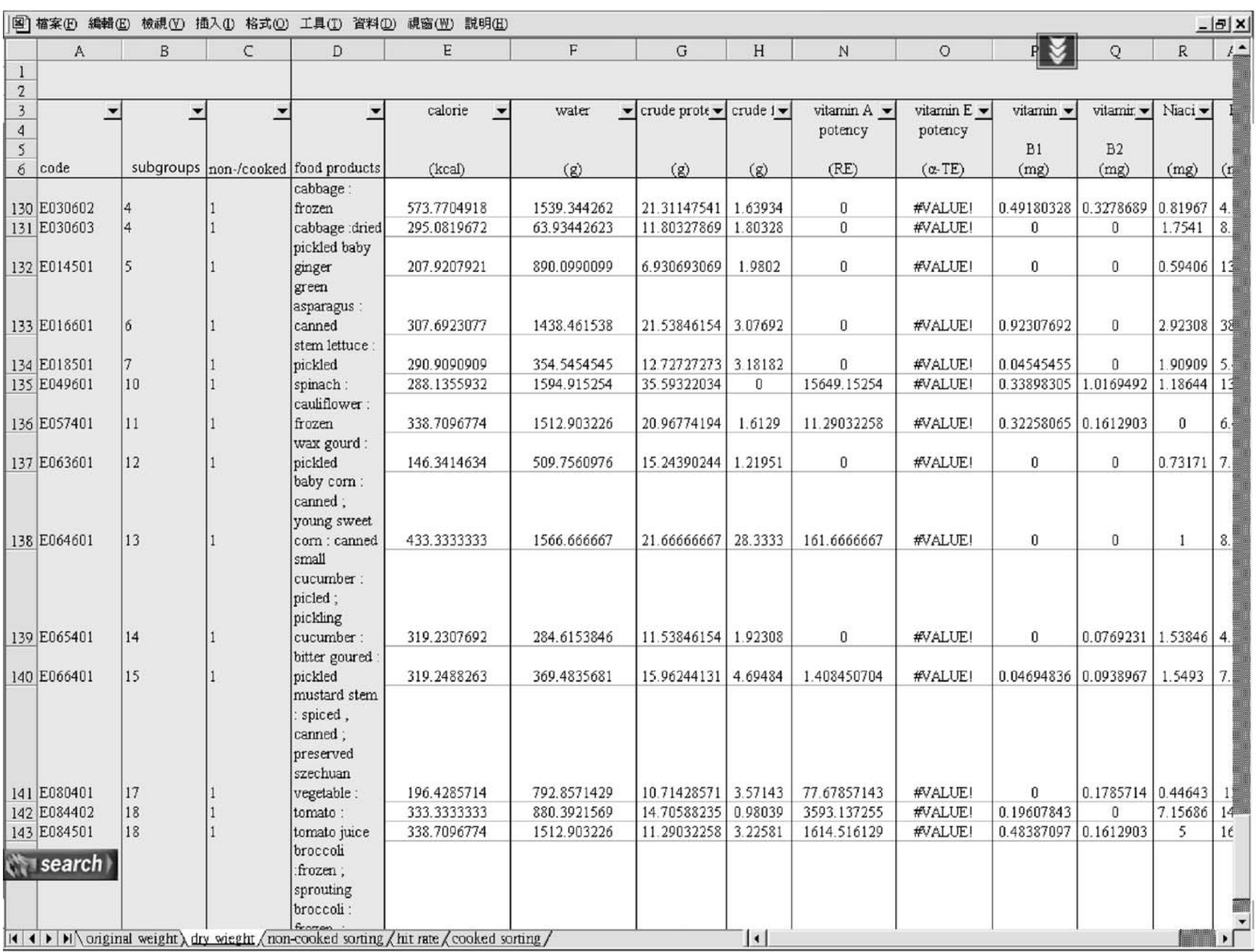
The 1226 foods from 18 groups were categorized into 239 subgroups based on their biological similarity by two nutritionists (Lee and Wu), as basic commodities or by way of recipe or food process involved in preparation. The Appendix to this paper provides the detailed food items of each VEGETABLES subgroup. Fig. 2 demonstrates the example of the VEGETABLES group. The cells "D130" and "D131" indicate frozen cabbage and dried cabbage are similar food items sorted into the 4th subgroup in cells of "B130" and "B131"; besides, they belong to processed methods of items coded as 1, which means processed food, in "C130" and "C131". There are 29 subgroups in column B coded as 0 (unsorted), 1 (carrots), 2 (radishes), ..., 28 (amaranths) and so on.

3.3. Calculation and ranking of CVs

The *CV*s of subgroups, except the unsorted ones, were calculated for both non-processed and processed foods (Fig. 3). Where all nutrient values are missing or undetectable they have been treated as zero for *CV*s.

3.4. Identification of unlikely values

The filter chosen comprised the combination of both the "*CV RANK*" of food and nutrient and also their "product" which leverages the "*hit*" and "*hit rate*" (Table 2). There were two possible approaches to set up the cut-offs for filter for the computerized process. The first approach was to choose a filter which could minimize both "*CV RANK*s" and their product. In the present exercise, we have adopted this method and found the most unlikely values for food varieties and only needed to review the optimized total *hits* for food groups and subgroups. In Table 2, among VEGETABLES, we observe that the 2 numbers for most *SHs* are 27 and 28; and that the "*SHR*" can reach more than 30%. We then deduce that the optimal



	A	B	C	D	E	F	G	H	N	O	P	Q	R	A
3–6	code	subgroups	non-/cooked	food products	calorie (kcal)	water (g)	crude prote (g)	crude f (g)	vitamin A potency (RE)	vitamin E potency (α-TE)	vitamin B1 (mg)	vitamir B2 (mg)	Niaci (mg)	(r
130	E030602	4	1	cabbage : frozen	573.7704918	1539.344262	21.31147541	1.63934	0	#VALUE!	0.49180328	0.3278689	0.81967	4.
131	E030603	4	1	cabbage :dried	295.0819672	63.93442623	11.80327869	1.80328	0	#VALUE!	0	0	1.7541	8.
132	E014501	5	1	pickled baby ginger	207.9207921	890.0990099	6.930693069	1.9802	0	#VALUE!	0	0	0.59406	13
133	E016601	6	1	green asparagus : canned	307.6923077	1438.461538	21.53846154	3.07692	0	#VALUE!	0.92307692	0	2.92308	38
134	E018501	7	1	stem lettuce : pickled	290.9090909	354.5454545	12.72727273	3.18182	0	#VALUE!	0.04545455	0	1.90909	5.
135	E049601	10	1	spinach :	288.1355932	1594.915254	35.59322034	0	15649.15254	#VALUE!	0.33898305	1.0169492	1.18644	13
136	E057401	11	1	cauliflower : frozen	338.7096774	1512.903226	20.96774194	1.6129	11.29032258	#VALUE!	0.32258065	0.1612903	0	6.
137	E063601	12	1	wax gourd : pickled	146.3414634	509.7560976	15.24390244	1.21951	0	#VALUE!	0	0	0.73171	7.
138	E064601	13	1	baby corn : canned ; young sweet corn : canned	433.3333333	1566.666667	21.66666667	28.3333	161.6666667	#VALUE!	0	0	1	8.
139	E065401	14	1	small cucumber : picled ; pickling cucumber :	319.2307692	284.6153846	11.53846154	1.92308	0	#VALUE!	0	0.0769231	1.53846	4.
140	E066401	15	1	bitter goured : pickled	319.2488263	369.4835681	15.96244131	4.69484	1.408450704	#VALUE!	0.04694836	0.0938967	1.5493	7.
141	E080401	17	1	mustard stem : spiced , canned ; preserved szechuan vegetable :	196.4285714	792.8571429	10.71428571	3.57143	77.67857143	#VALUE!	0	0.1785714	0.44643	1
142	E084402	18	1	tomato :	333.3333333	880.3921569	14.70588235	0.98039	3593.137255	#VALUE!	0.19607843	0	7.15686	14
143	E084501	18	1	tomato juice	338.7096774	1512.903226	11.29032258	3.22581	1614.516129	#VALUE!	0.48387097	0.1612903	5	16
				broccoli :frozen ; sprouting broccoli : frozen :										

Fig. 2. The Excel worksheet for the VEGETABLES group in the Taiwan Food Composition Tables. The columns are food item code (A), food subgroup (B), whether processed (0: unprocessed, 1: process): (C), food item name (D) and various nutrient values (E–X). All nutrients are expressed per 100 g dry weight. Where nutrient values are missing or undetectable they are represented by "0". All Chinese characters on the top of the frame are the menu (functional list) of Microsoft Excel.

	A	B	C	E	F	G	H	N	R	AA	AB	AC	AD	AE	AF
1	subgroups	name		Water	Crude protein	Crude fat	Carbohydrates	Vitamin E potenc	Vitamin B6	Zn		outliers	outliers		
2				2	3	4	5	11	15	24		1.5000	3.0000	1	
3	1	Carrot		*	*	*	*	*	*	*		#NUM!	#NUM!	#NUM!	#NUM!
4	2	Radishe		*	*	*	*	*	*	*		#NUM!	#NUM!	#NUM!	#NUM!
5	3	Bamboo Shoots		0.150606214	0.194315908	0.306018687	0.134899242	*	*	0.804570423		1.6435	2.5551	Vitamin A potenc	Vitamin B1
6	4	Cabbages		0.156017833	0.189787367	0.480040132	0.117868613	*	0.241244474	0.222795846		0.6964	1.0206	Vitamin A potenc	Vitamin B2
7	5	Ginger		0	1.414213562	0.282842712	0.094280904	*	*	0		0.6462	1.0265	Crude protein	Crude protein
8	6	Asparaguses		0.09114811	1.051788303	0.394960544	0.214055254	*	*	0.641946771		1.3776	2.0809	Crude protein	Vitamin A pot
9	7	Stems		0.439596331	0.514259477	0.346218258	0.275892966	*	*	0.346218258		1.1977	1.7659	Vitamin A poten	Vitamin B2
10	8	Chinese Chives		0.300608275	0.060268398	0.428847612	0.082208717	*	0.498972768	0.34015866		0.7290	1.0513	Fe	Ca
11	9	Bean Sprouts		0.192068286	0.401922923	0.065202215	0.4985612	*	*	0.187425894		2.0827	3.2379	Vitamin A potenc	Vitamin B2
12	10	Spinach		*	*	*	*	*	*	*		#NUM!	#NUM!	#NUM!	#NUM!
13	11	Cauliflower		*	*	*	*	*	*	*		#NUM!	#NUM!	#NUM!	#NUM!
14	12	Wax Gourd		*	*	*	*	*	*	*		#NUM!	#NUM!	#NUM!	#NUM!
15	13	Baby Corn		*	*	*	*	*	*	*		#NUM!	#NUM!	#NUM!	#NUM!
16	14	Cucumbers		0.06444648	0.2618914	0.341361894	0.155269738	*	0.061487546	0.061487546		0.8376	1.2646	Vitamin B1	P
17	15	Bitter Melons		0.711000626	0.265186994	0.538222049	0.162827705	*	0.075330995	0.163481722		1.4772	2.2433	Vitamin A potenc	Niacin
18	16	Sponge Gourds		0.030347931	0.103182964	0.255022118	0.008368128	*	0.185575535	0.582323232		1.5873	2.4461	Vitamin A potenc	Vitamin A pot
19	17	Mustard Stem Pickle		*	*	*	*	*	*	*		#NUM!	#NUM!	#NUM!	#NUM!
20	18	Tomato		*	*	*	*	*	*	*		#NUM!	#NUM!	#NUM!	#NUM!
21	19	Broccoli		*	*	*	*	*	*	*		#NUM!	#NUM!	#NUM!	#NUM!
22	20	Chinese Cabbages		0.593564491	0.590194984	1.001115005	0.426661972	*	0.760755665	0.709163551		0.9848	1.2350	Vitamin A poten	Vitamin B2
23	21	Celeries		0.317073027	0.670350273	1.046614016	0.760852418	*	0.300495905	0.58776247		1.7648	2.6434	Vitamin A potenc	Fe
24	22	Water Convolvuli		0.032370755	0.342738816	0.172465069	0.115209306	*	*	0.540193163		1.4249	2.2448	Vitamin B1	Vitamin A pot
25	23	Lettuces		0.24900722	0.341277759	0.226078518	0.388828285	*	0.680672087	0.127702221		0.8839	1.3185	Vitamin A poten	Vitamin B6
26	24	Bottle Gourds		0.029219288	0.129687707	0.686103609	0.044757822	*	0.446593757	0.027729678		1.0732	1.6998	Na	Vitamin A pot
27	25	Snake Gourds		0.333287368	0.099982195	0.173169008	0.079537388	*	1.024330124	0.310926386		0.8693	1.3013	Vitamin B6	Vitamin A pot
28	26	Sprouts		0.087201807	0.380549559	0.202365709	0.235718333	*	0.278278753	0.217829632		1.6538	2.5177	Vitamin A poten	Ca
29	27	Gymura		0.106244701	0.061686214	0.375485983	0.085946681	*	0.097838674	0.105231862		1.7563	2.7730	Vitamin B2	Niacin
30	28	Amaranths		0.481301566	0.196430228	0.471428623	0.13849031	*	0.885822737	0.157168558		1.5698	2.3782	Na	Vitamin B2
31															
32															
33	outliers	1.5000	x	0.6992	0.9056	0.8285	0.5043	#NUM!	1.4233	1.1895					
34	outliers	3.0000	x	1.0732	1.3532	1.1813	0.7529	#NUM!	2.1660	1.8177					
35	1			Bitter Melons	Ginger	Celeries	Celeries	#NUM!	Snake Gourds	Bamboo Shoots					
36	2			Chinese Cabbages	Asparaguses	Chinese Cabbag	Bean Sprouts	#NUM!	Amaranths	Chinese Cabbages					
37	3			Amaranths	Celeries	Bottle Gourds	Chinese Cabbages	#NUM!	Chinese Cabbages	Asparaguses					
38	4			Stems	Chinese Cabbages	Bitter Melons	Lettuces	#NUM!	Lettuces	Celeries					
39	5			Snake Gourds	Stems	Cabbages	Stems	#NUM!	Chinese Chives	Sponge Gourds					
40	6			Celeries	Bean Sprouts	Amaranths	Sprouts	#NUM!	Bottle Gourds	Water Convolvuli					
41	7			Chinese Chives	Sprouts	Chinese Chives	Asparaguses	#NUM!	Celeries	Stems					
42	8			Lettuces	Water Convolvuli	Asparaguses	Bitter Melons	#NUM!	Sprouts	Chinese Chives					
43	9			Bean Sprouts	Lettuces	Gymura	Cucumbers	#NUM!	Cabbages	Snake Gourds					
44	10			Cabbages	Bitter Melons	Stems	Amaranths	#NUM!	Sponge Gourds	Cabbages					
45	11			Bamboo Shoots	Cucumbers	Cucumbers	Bamboo Shoots	#NUM!	Gymura	Sprouts					
46	12			Gymura	Amaranths	Bamboo Shoots	Cabbages	#NUM!	Bitter Melons	Bean Sprouts					
47	13			Asparaguses	Bamboo Shoots	Ginger	Water Convolvuli	#NUM!	Cucumbers	Bitter Melons					
48	14			Sprouts	Cabbages	Sponge Gourds	Ginger	#NUM!	#NUM!	Amaranths					
49	15			Cucumbers	Bottle Gourds	Lettuces	Gymura	#NUM!	#NUM!	Lettuces					
50	16			Water Convolvuli	Sponge Gourds	Sprouts	Chinese Chives	#NUM!	#NUM!	Gymura					
51	17			Sponge Gourds	Snake Gourds	Snake Gourds	Snake Gourds	#NUM!	#NUM!	Cucumbers					
				Bottle Gourds	Gymura	Water Convolvuli	Bottle Gourds	#NUM!	#NUM!	Bottle Gourds					
				Ginger	Chinese Chives	Bean Sprouts	Sponge Gourds	#NUM!	#NUM!	Ginger					
54	20			#NUM!	#NUM!	#NUM!	#NUM!	#NUM!	#NUM!	#NUM!					
55	21			#NUM!	#NUM!	#NUM!	#NUM!	#NUM!	#NUM!	#NUM!					
56	22			#NUM!	#NUM!	#NUM!	#NUM!	#NUM!	#NUM!	#NUM!					

search

original weight / dry wieght / non-cooked sorting / hit rate / cooked sorting

Fig. 3. The coefficients of variation (*CV*s) for the subgroups of the VEGETABLES group in the Taiwan Food Composition Tables. For the purpose of overview, this worksheet has omitted columns, such as D, I, J, etc... Values in cells E3 to AA30 are the CVs of subgroups. "*" means non-calculable CV. All CVs were ranked in descending order for food subgroups (from the column "AE" and beyond) and for each nutrient (from the 35th row and below). Two cut-offs for outliers have been defined as mean +/- 1.5 and 3.0 times standard deviation calculated (33rd and 34th row for nutrients and "AC" and "AD" for subgroup). "#NUM!" indicates "un-rankable". Highlight in this worksheet are for Panel discussion. All Chinese characters on the top of the frame are the menu (functional list) of Microsoft Excel.

parameters of "rank" and "product" for the filters are 2 and 20, respectively, in Table 2.

The second approach was to define a certain "*SHR*" based on resources and capacity to prioritize unlikely values. To choose a smaller "*CV RANK*" and CRP would result higher "*SHR*", which would mean more unlikely values, and vice versa.

As shown in Fig. 3, we first ranked the food subgroups *CV*s for each nutrient in descending order (from the 35th row and below). Then, we ranked the nutrient *CV*s for each food subgroup (from the column "AE" and beyond). Two cut-offs for outliers have been defined as mean $\pm$ 1.5 and 3.0 times the standard deviation calculated, respectively (33rd and 34th row for nutrients and "AC" and "AD" for subgroups). In Fig. 4, each "*hit*" is shown as the two components of a vector (x,y) in parenthesis, with the food subgroup *CV RANK* as (x) and the nutrient type *CV RANK* as (y), and the product (CRP) outside the parenthesis. Accordingly, the worksheet "*hit rate*" is represented as a vector (two *CV RANK*s, one for food subgroup and one for nutrient, respectively, multiplied). The numeric product of the elements of the vector is shown as in "E17" by "(1,6) = 6".

For the VEGETABLES group, a total of 82 *hits* were screened by the computerized process. A *hit rate* (*HR*) was calculated as (sum of *hits*/total piece of information) $\times$ 100%, where total pieces of information are the numbers of foods multiplied by the number of nutrients. The *HR* for the VEGETABLES group is 2.48%, which is calculated as [82 *hits*/(138 items $\times$ 24 nutrients)] $\times$ 100%. This computer process can only assess a food subgroup which comprises two or more foods. Some performance parameters have been shown in the collective food group worksheets for "*hits*", as in Fig. 4 for the VEGETABLES group. Table 1 lists sum of *hits* and *HR* for all 18 food groups.

3.5. Judgment by a food and nutrition expert panel

The Expert Panel was able to review each *hit* of a screened subgroup or nutrient for unlikely values. Suggestions made for further investigation were "recheck in laboratory", "definitely describes food item/species", "too high/low nutrient content" and so on.

"*Satisfied hits*" (*SH*) were system-defined hits which the Expert Panel agreed were legitimate errors worthy of further evaluation, as highlighted in Fig. 4. The satisfied *hit rate* (*SHR*) was calculated as (*Satisfied hits*/sum of *hits*) $\times$ 100%. "*Missing hits*" (*MH*) were those identified by the Expert Panel to avoid false negatives generated by the third and fourth steps, but not by the computerized process. In this example, the cell "C31" in Fig. 4 displayed the sum of *hits* as 82, cell "C32" displayed *SH* as 27, cell "C34" displayed the few MH (e.g. "****" on cell I22) as 6, and cell "C33" displayed the *SHR* without MH as 32.93%.

In all cases the Expert Panel took into account FCT from not only Taiwan, but also China (China Institute of Nutrition and Food

檔案(F) 編輯(E) 檢視(V) 插入(I) 格式(O) 工具(T) 資料(D) 視窗(W) 說明(H)

	A	B	C	E	F	G	H	I	M	O	P	Q	S	T	U	W	X	Y	Z	AA
1	subgroups	name		Water	Crude protein	Crude fat	Carbohydrate	Crude fiber	Vitamin A p	Vitamin B1	Vitamin B2	Niacin	Vitamin B12	Vitamin C	Na	Ca	Mg	P	Fe	Zn
2				2	3	4	5	6	10	12	13	14	16	17	18	20	21	22	23	24
3	1	Carrot																		
4	2	Radishe																		
5	3	Bamboo Shoots							(10,1)=10	(5,2)=10				(3,3)=9		(3,4)=12				(1,5)=5
6	4	Cabbages				(5,4)=20			(2,1)=2		(10,2)=20									
7	5	Ginger			(1,1)=1				(3,1)=3			(4,4)=16		(5,3)=15						
8	6	Asparaguses			(2,1)=2				(14,2)=28										(5,3)=15	
9	7	Stems							(6,1)=6		(3,2)=6			(2,3)=6	(5,4)=20			(2,6)=12		
10	8	Chinese Chives														(7,2)=14			(7,1)=7	
11	9	Bean Sprouts					(2,8)=16		(3,1)=3		(1,2)=2			(1,3)=3	(4,4)=16	(2,5)=10	****	****	****	****
12	10	Spinach																		
13	11	Cauliflower																		
14	12	Wax Gourd																		
15	13	Baby Corn																		
16	14	Cucumbers								(1,1)=1								(1,2)=2		
17	15	Bitter Melons		(1,6)=6					(9,1)=9	(4,3)=12		(1,2)=2			****	(4,5)=20			(3,4)=12	
18	16	Sponge Gourds						(3,5)=15	(3,1)=3		(5,4)=20				(1,1)=1				(1,3)=3	
19	17	Mustard Stem Pickle																		
20	18	Tomato																		
21	19	Broccoli																		
22	20	Chinese Cabbages		(2,13)=26		(2,4)=8		****	(8,1)=8	(3,3)=9	(4,2)=8		****			****			****	(2,7)=14
23	21	Celeries				(1,3)=3	(1,8)=8	(1,5)=5	(7,1)=7								(1,6)=6		(2,2)=4	
24	22	Water Convolvuli							(11,2)=22	(2,1)=2				(4,3)=12			(2,4)=8			
25	23	Lettuces							(13,1)=13											
26	24	Bottle Gourds				(3,4)=12			(12,2)=24						(1,1)=1					
27	25	Snake Gourds							(16,2)=32						(6,3)=18					
28	26	Sprouts							(1,1)=1	(6,3)=18		(3,5)=15				(1,2)=2				
29	27	Gynura									(6,1)=6	(2,2)=4							(4,5)=20	
30	28	Amaranths						(2,5)=10			(2,2)=4				(3,1)=3					
31		sum of hits:	82	2	2	4	2	3	15	6	7	4	0	5	6	5	2	2	6	2
32		satisfied hits:	27																	
33		SHR	32.93%																	
34		missing hits:	6																	
35																				

Fig. 4. The ''satisfied hit rate'' (*SHR*) derived from food subgroups and nutrient type for the VEGETABLES group of the Taiwan Food Composition Tables. For the purpose of overview, this worksheet has omitted columns J, K, and L. Satisfied hits (Row 32) were system-defined hits which the Expert Panel agreed were legitimate errors worthy of further evaluation but without missing hits. SHR (Row 33) is the satisfied hit rate (C33 = C32/C31) without missing hits. The Panel detected unlikely values which are shown as ''****''. Highlights in this worksheet are Panel agreed. All Chinese characters on the top of the frame are the menu (functional list) of Microsoft Excel.

Safety, 2002) and Japan (Kagawa, 1999) in their review and evaluation of ''*hits*'' (see below).

3.6. Methodological gains

The audit process screened for potentially unlikely values in 18 food groups, by way of 239 subgroups of 1226 food items and 24 nutrients in the Taiwan FCT and the Expert Panel reviewed the values to determine likely validity or need for revision. Table 1 is a summary of the performance of the process. This system detected 777 (759 from the computerized process and 18 from the Panel process) unlikely nutrient values in the FCT (comprising 29,424 (1226 items $\times$ 24 nutrients) pieces of information). For this FCT, the *HR* (or probable error rate, *PER*) from the auditing process was 2.6% (777/29,424). The process alone performed nearly 38 times (29424/777) better than the manual only method. Even though there were 18 missing nutrient errors identified by the Panel, and not by the computerized process, they were readily identifiable in the subgroups screened by changing the CRP from 20 to 30 (Table 1). However, in this case, the *HR* would decrease due to a false positive as indicated in Table 2.

The Expert Panel reviewed values in relation to the FCT of Taiwan, China and Japan and recommendations were made to regard 116 items as genuinely unlikely (*SH*) and to require further laboratory analysis, so that the FCT could be revised. The overall *SHR* was 14.9% (ranged 1.4–37.6) for the various food groups.

3.7. Accuracy and applicability of the method

There were high *SHR* in several food groups, 19.3%, 32.0%, 30.7%, and 37.6% in the NUTS & SEEDS, FRUITS, VEGETABLES, and MEATS groups, respectively (Table 1). In food groups with many processed food items, few unlikely nutrient values were assigned by the Panel; these were the CEREALS, PULSES, EGGS, MILKS, SEASONINGS & SPICES, CONFECTIONERIES, and PROCESSED FOODS groups with *SHR* ranging from 1.4% to 9.0%.

Some food items, mostly processed and treated, were not well subgrouped for reliable assessment of variation: they had no *hit* in this system. These were the POTATOES & STARCHES, ALGAE, MUSHROOMS, FISH AND SEAFOOD, FATS & OILS, SUGARS & SWEETENERS, and BEVERAGES groups. For example, many shrimps and fish items were not generic and could not be subgrouped reliably. This also applied to foods that had been used to prepare other items as with certain fats, oils, sugars, sweeteners, potatoes, and starches. Additionally, foods of low nutrient density like FATS & OILS and SUGARS & SWEETENERS have few nutrients to detect, except for fat-soluble vitamins A and E in FATS & OILS; there were no *CV*s or vectors to calculate or assess. For ALGAE, there were only 5 items and too few to create subgroups. For MUSHROOMS, even though there were 6 subgroups, there were too few food items to allocate subgroups for reliable variation.

3.8. Methodological limitations

The *hits* which satisfied the Panel as legitimate errors were greater when there were enough food items that could be correctly assigned to a subgroup, especially when there were ''unprocessed foods''. This data audit performed better for unprocessed foods, which could be confidently subgrouped.

In terms of number of food items needed in a subgroup, theoretically, a population *CV* can be calculated with an *n* of 2. Although the meaning and utility of the *CV*s from small subgroups may be questioned, we have demonstrated that this approach still yields useful information. For example, in the bean sprouts subgroup, with 2 items only, we have identified that mung bean sprouts, but not soy bean sprouts, have an unlikely high vitamin C content.

For calculating *CV*s, we have had to treat all missing information as undetectable values and zero. This system cannot distinguish between true zero and missing information. Foods of low nutrient density (e.g. SUGAR) have too few nutrients to calculate *CV*s.

3.9. External validity

To exemplify, the FRUITS in the Taiwan FCT for the subgroups GRAPES and WATERMELONS are shown in Table 3. They are juxtapositioned alongside those from the Chinese and Japanese FCT (China Institute of Nutrition and Food Safety, 2002; Kagawa, 1999) to assist in the recognition of unlikely values. In the GRAPES subgroup of the Taiwan FCT, there were 3 items (California grape (D019400), White grape (D020400) and Grape (D021400)) which

Table 3
The subgroups "grapes" and "watermelons" in the FRUITS group of Taiwan Food Composition Tables: a comparison with Chinese and Japanese data for vitamin A activity.

Code	Food items	Suggestions	Vitamin A activity (RE)
Subgroup of Grapes			
D019400	California grape	- Indicate plant breed	37.5
D020400	White grape	- Too variable for vitamin A activity	1410
D021400	Grape		0
Chinese FCT	Grape		70.8
	Hongmeigui[#]		0
	Kyoho		38.5
	Menindee		76.9
	Muscat		22.9
	Dark-skinned		86.2
Japanese FCT	Grape		32.1
Subgroup of Watermelons			
D026400	Xiaoyu[#]; yellow pulp	- Too variable for vitamin A activity	3.33
D027400	Watermelon; red pulp		1810
Chinese FCT	Watermelon		1119
	Jinxin[#] 1		148
	Zhenzhou[#] 3		530
	Zhongyu[#] 6		494
Japanese FCT	Watermelon		707

[#] Phonetically translated on the basis of Mandarin pronunciation.

the Panel regarded as "too variable in vitamin A potency" when considered in conjunction with similar food items for China and Japan (China Institute of Nutrition and Food Safety, 2002; Kagawa, 1999). It was also concluded that: "The analytical laboratory should indicate explicitly what cultivar the D019400 is, by its scientific name (*Vitis spp.* L.) as in the Japanese FCT (Japan Science and Technology Agency, 2005) or by description as in the Encyclopaedia Britannica and Wikipedia (Wikipedia, 2006)."

Since the D020400 (White grape) contained 1410 RE vitamin A activity per 100 g DW, markedly greater than the ones (0, 22.9, 38.5, 70.8, 76.9, and 86.2 RE) from the China FCT and the one (32.1 RE) from the Japan FCT. The vitamin A activity of the subgroup GRAPES was unlikely.

Likewise, in the WATERMELONS subgroup (D027400, Watermelon) contained 1810 RE vitamin A activity per 100 g DW that was much greater than the ones (148, 494, 530, and 1119 RE) from the China FCT and the one (707 RE) from the Japan FCT. This meant that vitamin A activity in the WATERMELONS subgroup was unlikely.

4. Discussion

We present an informatics process to identify extreme nutrient values for foods in the Taiwan FCT from a data audit of coefficients of variation (*CV*s). These included minerals among NUTS & SEEDS, vitamin A among FRUITS, crude fat and carbohydrate and vitamins among VEGETABLES, and all nutrients among MEATS. This is in close agreement with the experience of nutritionists and suggests that there are needs for laboratory re-checks and revision of various food composition tables.

Nutritionists often recommend intakes from particular food subgroups in order to achieve specific nutrient adequacy. Citrus fruits are considered to be vitamin C rich fruits; pro-vitamin A is abundant in dark green or yellow fruits and vegetables. Intuitively, this informatics process for data quality audit of the FCT is based on the idea that similar foods should consist of similar levels of nutrients; this is thought to agree with common sense and current theory which underpins dietary guidelines and food guides. At the same time, the evaluation of nutrient variety represented by a diversity of food groups, with defined nutritional characteristics, can assist with the universal dietary guideline that encourages variety and for which there is growing evidence (Wahlqvist and Lee, 2007).

We used *CV*s for the outliers, being those ranked highest either for food subgroups or for particular nutrients, as well as the least products of these 2 ranks to find the most suspect pieces of nutrient information – they were marked as unlikely values and constituted "*hits*". This identification of unlikely values acted as a filter to evaluate the entire food composition table. The ranks for *CV*s (*CV RANK*) and their product (CRP) could provide a threshold and minimize the need for manual review of each piece of nutrient information.

The filtered and unlikely nutrient value was far from the expected, if similar foods have like levels of nutrients. The Panel reviewed and discussed the unlikely values in the Taiwan FCT as "system-defined hits" (*SH*) and also considered them in conjunction with the China and Japan Food Composition Tables (China Institute of Nutrition and Food Safety, 2002; Kagawa, 1999). Several extreme food nutrient values were identified. This made it possible to reappraise them and to comment about the appropriateness of analytical techniques for some foods.

The variability of analytical data in FCT is high for micronutrients (Leclercq et al., 2001; Salvini, 1997; Stryker et al., 1991) and is moreover increased by taking other factors into consideration: sources of plants, animal husbandry, storage, transport and marketing. Food processing methods vary by ingredient content and composition with changes in formulation and production. There are also problems from artefacts introduced though inadequacy of sampling, differences in analytical methods, lack of use of quality assurance programs, and differences in deriving protein and other macronutrient values with the compilation of FCT (Leclercq et al., 2001). The inappropriate compilation of FCT has caused gross errors in the assessment of nutrient intake levels in European countries. These have included time trends and between-country estimates for food and nutrient intakes. Improper food composition data limit the understanding of the relationships among nutrients, health and disease (Dwyer, 1994).

Our informatics approach, with more efficient use of an Expert Panel, has provided a way to check whether data might derive from incompatible variation. "*Hits*" for extreme variation of nutrient values have provided an audit for FCT. The approach has enabled the compilation of more correct sources of data for FCT by verifying the definitions, analytical methods and modes of expression originally used. At the same time, where the intra-subgroup variation is real, the approach may provide opportunities to refine nutritional advice, based on reliable FCT.

The Panel was convened to clarify food similarity when audit process *hits* were discovered. The clearer and more certain identity of food items, by scientific names and illustration, is recommended for FCT (Poortvliet et al., 1992; Puwastien, 2002; Rand et al., 1991; Southgate and Greenfield, 1992). Scientific names would enable foods within a subgroup to be more biologically and nutritionally similar and would allow a more convincing subgroup to be created in the first place. This process would be strengthened by having foods labelled in accordance with a scientific classification, as well as secular names, as has been the objective in the INFOODs project through the FAO, United Nations University and IUNS (International Union of Nutritional Sciences). Illustrations (photos) of food items, appropriately referenced, are also recommended. INFOODs recommends constructing a complete internet homepage of foods throughout the world to assist when food items are difficult to identify and name (Masson, 1999; Rand and Young, 1984). Food illustrations, the various secular, local or traditional food names, scientific names and adequate food descriptions would help practitioners with their dietary assessments when utilizing FCT and help consumers and manufacturers to implement recommendations, especially on Internet-based food information.

Another challenge in the assessment of the utility and validity of FCT is the overall food culture (whether by geography or ethnicity), which may have more or fewer food items in a particular food group or subgroup. Examples would be the few or many types of cheese in European subgroups of the dairy food group. Another would be the spices group which could comprise few or many items depending on the food culture in question, and which could lead to an enormous potential variability in nutrient contents.

In the case of the spices group, a particular difficulty arises: this is partly because of the small quantities that may have disproportionate biological effects, partly because the bio-active food components are likely to extend to many different phytonutrients, and partly because of different potential effects dependent on background diet. This example typifies how the demands for an "informatics process" applied to FCT are likely to grow rapidly beyond the present requirements for macro- and micronutrients.

The "informatics" approach in the present paper acknowledges the possibility that an outlier nutrient content in a food subgroup may not be the same from one food culture to another. This may be because the range of food species of apparently biologically similar foods may vary considerably. In turn, the median nutrient value and the *CV*s may be skewed. However, there is no reason not to test different cut-off points, FCT by FCT, and make the criteria known.

From a practical point of view, where those responsible for FCT accuracy in various settings wish to use our method, there will be translational requirements. One has to do with differences in numbers of food categories and items and in number and type of nutrients. The cut-off points (*CV RANKS* and CRP) might be affected by the number of nutrients in a database because the variations and precisions of macro- and micronutrients are influential in laboratory processes and methods.

In conclusion, an informatics process for data audit is able to identify potentially unlikely nutrient values in apparently similar foods. It can help improve the accuracy and utility of FCT in the Taiwan area. It should also improve FCT in other countries and be responsive to new demands on FCT as long as they become more and more digitized. The demands on quality assurance (QA) programs for FCT as databases expand, with growing numbers of foods and bio-active food components subject to measurement

Appendix A

Subgroup assignments for the VEGETABLES group.[a]

Code	Subgroup	Name (Food code)[b]
0	Unsorted	(Un-pro[c]): Burdock (E001400), Lotus root (E003400), Dasheen stalk (E007400), Chinese water-chestnut (E010400), Green asparagus juice (E016502), Shallot (E020400), Onion (E023400), Scallion (E024400), Leek (E025400), Pakchoi (E026400), Basil (E027400), Chinese knotweed (E029400), Sweet potato leaves (E031400), Leaf mustard (E034400), Coriander (E036400), Edible rape (E038400), Ching- chiang pakchio (E039400), Garland chrysanthemum (E042400), Shannon coriander (E043400), Crown green (E045400), Fern bud (E053400), Solanum nigrum (E054400), Chayote stem shoot (E055400), Day-lily flower (E060400), Vegetable sponge flower (E061400), Rape flower (E062400), Eggplant (E067400), Pumpkin (E069400), Bell pepper (E070400), Chayote stem gourd (E071400), Yellow okra (E072400), Chili (E076400), Mountain potato (E086500), Yam (E088600), White mugwort (E091600), Green water cress (E092600), Feather cockscomb E094600, Nodding burnweed (E095600), Red stem pearl vegetable (E096600), Fresh Mastixia (E097600), Cedrus (E098600), Coral flower (E101600), Houttuynia (E105600), Guogou vegetable fern, (E106600), Garbo melon (E107600), Dragon grass (E111600), Mint (E112600), Chinese onion (E113600), Notoginseng (E114600), Aloe (E115600); (Pro[d]): Chili sauce (E076401), Pickled mustard leaves (E078400), Pickled leaf mustard (E079400), Oriental pickled melon (E081400), Pickled mustard stem (E082400), Pickled cucumber in soy sauce (E083300)
1	Carrots	(Un-pro):Carrot (E002400); (Pro): Frozen carrot (E002601)
2	Radishes	(Un-pro): Radish (E004400); (Pro): Pickled radish (E004501), Dried radish (E004502)
3	Bamboo shoots	(Un-pro): Bamboo shoot (E005400) Betel-nut palm stem (E006400) Water-bamboo (E009400) Ma bamboo shoot (E012400); (Pro): Dried bamboo shoots (E005503) Frozen bamboo shoot (E005601) Canned bamboo shoot (E005602)
4	Cabbages	(Un-pro): Kohlrabi (E011400), Cabbage (E030400), Cabbage (E030401), Purple cabbage (E050400), Baby cabbage (E100600); (Pro): Frozen cabbage (E030602), Dried cabbage (E030603)
5	Gingers	(Un-pro): Young ginger (E014400), Ginger (E017400); (Pro): Pickled young ginger (E014501)
6	Asparagus	(Un-pro): Green asparagus (E016400), Asparagus (E019400); (Pro): Canned green asparagus (E016601)
7	Stem lettuce	(Un-pro): Stem lettuce (E018400), Snow pea stem (E047400); (Pro): Pickled stem lettuce (E018501)
8	Chinese chives	Chinese chives (E021400), Yellow Chinese chives (E022400), Chinese leek flowerbud (E059400)
9	Bean sprouts	Soy bean sprouts (E013400), Mung bean sprouts (E015400)

Appendix A (*Continued*)

Code	Subgroup	Name (Food code)[b]
10	Spinaches	(Un-pro): Spinach (E049400); (Pro): Frozen spinach (E049601)
11	Cauliflower	(Un-pro): Cauliflower (E057400); (Pro): Frozen cauliflower (E057401)
12	Wax gourd	(Un-pro): Wax gourd (E063400); (Pro): Pickled wax gourd (E063601)
13	Baby corns	(Un-pro): Baby corn (E064400);(Pro): Canned baby corn (E064601)
14	Cucumbers	(Un-pro): Small cucumber (E065400), Cucumber (E068400); (Pro): Pickled cucumber (E065401)
15	Bitter gourds	Un-pro): Bitter gourd (E066400), Wild bitter gourd (E102600), Young wild bitter gourd (E103600); (Pro): Pickled bitter gourd (E066401)
16	Luffa	Luffa (E073400), Angled luffa (E077400)
17	Mustard tuber	(Pro): Canned hot pickled mustard tuber (E080401), Hot pickled mustard tuber (E080500)
18	Tomatoes	(Un-pro): Tomato (E084400); (Pro): Canned tomato (E084402), Tomato juice (E084501)
19	Broccoli	(Un-pro): Broccoli (E058400); (Pro): Frozen broccoli (E058601)
20	Chinese cabbages	San-tong cabbage (E028400), Chinese cabbage (E032400), Cuiyu cabbage (E109600), Baby Cuiyu cabbage (E110600)
21	Celery	Chinese celery (E033400), Celery (E040400), Hill celery (E087600)
22	Water convolvulus	Water convolvulus (E037400), Water cress (E041400)
23	Lettuce	Head lettuce (E051400), Lettuce leaves (E052400), Leaf lettuce (E056400)
24	Bottle gourd	Bottle gourd (E074400), Bottle gourd (E075400)
25	Snake gourd	White snake gourd (E089600), Green snake gourd (E108600)
26	Sprouts	Alfalfa sprouts (E008400), Chinese kale (E035400), Chinese kale sprouts (E093600)
27	Gynura	Red gynura (E044400), White gynura (E090600)
28	Amaranth	Red amaranth (E046400), Amaranth (E048400), Purslane(E099600), Wild amaranth (E104600)

and nutritional evaluation, will be great. The present approach should allow QA commensurate with this trend.

Acknowledgements

This study was supported by a grant (DOH94-TD-F-113-067-(2)) from the Research Fund of Bureau of Food Sanitation, Department of Health, Taiwan, Republic of China.

References

Alexander, J.T., Cheung, W.K., Dietz, C.B., Leibowitz, S.F., 1993. Meal patterns and macronutrient intake after peripheral and PVN injections of the alpha 2-receptor antagonist idazoxan. Physiol. Behav. 53, 623–630.

Bowers, L., 2002. An audit of referrals of children with autistic spectrum disorder to the dietetic service. J. Hum. Nutr. Diet. 15, 141–144.

Buzzard, I.M., Asp, E.H., Chlebowski, R.T., Boyar, A.P., Jeffery, R.W., Nixon, D.W., Blackburn, G.L., Jochimsen, P.R., Scanlon, E.F., Insull, W.J.R., et al., 1990. Diet intervention methods to reduce fat intake: nutrient and food group composition of self-selected low-fat diets. J. Am. Diet. Assoc. 90 (42-50), 53.

China Institute of Nutrition and Food Safety, 2002. Chinese Food Composition. China CDC, Beijing (in Chinese).

Dartois, A.M., Ducamp, S., Decaux, F., Broyer, M., 1989. Computer-assisted diet therapy in pediatric kidney diseases. Pediatrie 44, 197–202 (in French).

Deharveng, G., Charrondiere, U.R., Slimani, N., Southgate, D.A., Riboli, E., 1999. Comparison of nutrients in the food composition tables available in the nine European countries participating in EPIC. European Prospective Investigation into Cancer and Nutrition. Eur. J. Clin. Nutr. 53, 60–79.

Dwyer, J.T., 1994. Future directions in food composition studies. J. Nutr. 124, 1783S–1788S.

Fabiansson, S.U., 2006. Precision in nutritional information declarations on food labels in Australia. Asia Pac. J. Clin. Nutr. 15, 451–458.

Farran-Codina, A., Boatella-Riera, J., Serra-Majem, L., Ribas, L., Raffcas-Martinez, M., Codony-Salcedo, R., 1994. General criteria for the development and use of data tables and systems on food composition. Rev. Sanid. Hig. Publica. (Madr) 68, 427–441.

Fidanza, F., Perriello, G., 2002. Validation of the Italian food composition database of the European institute of oncology. Eur. J. Clin. Nutr. 56, 1004–1010.

Guilland, J.C., Aubert, R., Lhuissier, M., Peres, G., Montagnon, B., Fuchs, F., Merlet, N., Astorg, P.O., 1993. Computerized analysis of food records: role of coding and food composition database. Eur. J. Clin. Nutr. 47, 445–453.

Institute for Interlaboratory Studies, 2006. Statistics in Porficiency Tests, Vol. 2006.

ISO/IEC, 1997a. ISO/IEC Guide 43-1—Part 1: Development and operation of proficiency testing schemes. In: Proficiency testing by interlaboratory comparisons. ISO/IEC.

ISO/IEC, 1997b. ISO/IEC Guide 43-2—Part 2: Selection and use of proficiency testing schemes by laboratory accreditation bodies. In: Proficiency testing by interlaboratory comparisons. ISO/IEC.

Japan Science and Technology Agency, 2005. Food Composition Database (in Japanese).

Kagawa, Y. (Ed.), 1999. Standard Tables of Food Composition in Japan. Women's Nutrition College Publishing Department, Tokyo (in Japanese).

Leclercq, C., Valsta, L.M., Turrini, A., 2001. Food composition issues—implications for the development of food-based dietary guidelines. Public Health Nutr. 4, 677–682.

Masson, L., 1999. LATINFOODS and its role in the generation and compilation of data for Latin America. Arch Latinoam Nutr. 49, 89S–91S (in Spanish).

Matsuda-Inoguchi, N., Shimbo, S., Nakatsuka, H., Watanabe, T., Higashikawa, K., Ikeda, M., 2004. Effects of revision of Japanese food composition tables on estimation of nutrient intakes, with reference to age-dependent differences. Public Health Nutr. 7, 901–909.

Mccullough, M.L., Karanja, N.M., Lin, P.H., Obarzanek, E., Phillips, K.M., Laws, R.L., Vollmer, W.M., O'Connor, E.A., Champagne, C.M., Windhauser, M.M., 1999. Comparison of 4 nutrient databases with chemical composition data from the Dietary Approaches to Stop Hypertension trial. DASH Collaborative Research Group. J. Am. Diet. Assoc. 99, S45–S53.

NATA, 1996. New Statistics for Proficiency Testing Programs. National Association of Testing Athorities (NATA), Australia.

Phillips, K.M., Patterson, K.Y., Rasor, A.S., Exler, J., Haytowitz, D.B., Holden, J.M., Pehrsson, P.R., 2006. Quality-control materials in the USDA National Food and Nutrient Analysis Program (NFNAP). Anal. Bioanal. Chem. 384, 1341–1355.

Poortvliet, E.J., Klensin, J.C., Kohlmeier, L., 1992. Rationale document for the Eurocode 2 food coding system (version 91/2). Eur. J. Clin. Nutr. 46 (Suppl. 5), S9–S24.

Probst, Y.C., Krnavek, C., Lockyer, L., Tapsell, L.C., 2004. Developing a self-administered computer assisted dietary assessment tool for use in primary healthcare practice: perceptions of nutrition and computers in older adults with T2DM. Asia Pac. J. Clin. Nutr. 13, S136.

Puwastien, P., 2002. Issues in the development and use of food composition databases. Public Health Nutr. 5, 991–999.

Rand, W.H., 1985. Food composition data: problems and plans. J. Am. Diet. Assoc. 85, 1081–1083.

Rand, W.M., 1987. INFOODS and food composition data. Arch. Latinoam. Nutr. 37, 609–617 (in Spanish).

Rand, W.M., Pennington, J.A.T., Murphy, S.P., Klensin, J.C., 1991. Compiling Data for Food Composition Data Bases. The United Nations University, Tokyo, Japan, p. 77.

Rand, W.M., Young, V.R., 1984. Report of a planning conference concerning an international network of food data systems (INFOODS). Am. J. Clin. Nutr. 39, 144–151.

Rodriguez, M.C., Coupeau, I., Larralde, J., Martinez, A., 1995. Development of a computer program to assess nutritional status of infants and preschool children. Arch. Latinoam. Nutr. 45, 274–280 (in Spanish).

Salvini, S., 1997. A food composition database for epidemiological studies in Italy. Cancer Lett. 114, 299–300.
Samsonov, M.A., Meshcheriakova, V.A., Shadur, S.S., Shekhter, E.G., 1986. A computer-assisted method of evaluating the actual nutrition of patients with ischemic heart disease during their clinical observation. Vopr. Pitan. 18–21 (in Russian).
Sasaki, S., Kobayashi, M., Tsugane, S., 1999. Development of substituted fatty acid food composition table for the use in nutritional epidemiologic studies for Japanese populations: its methodological backgrounds and the evaluation. J. Epidemiol. 9, 190–207.
Shai, I., Vardi, H., Shahar, D.R., Azrad, A.B., Fraser, D., 2003. Adaptation of international nutrition databases and data-entry system tools to a specific population. Public Health Nutr. 6, 401–406.
Sharp, M.M., Ahmed, K., 1983. A computer application for dietary analysis in clinical nutrition. J. Can. Diet. Assoc. 44, 228–234.
Shimbo, S., Hayase, A., Murakami, M., Hatai, I., Higashikawa, K., Moon, C.S., Zhang, Z.W., Watanabe, T., Iguchi, H., Ikeda, M., 1996. Use of a food composition database to estimate daily dietary intake of nutrient or trace elements in Japan, with reference to its limitation. Food Addit. Contam. 13, 775–786.
Southgate, D.A., Greenfield, H., 1992. Principles for the preparation of nutritional data bases and food composition tables. World Rev. Nutr. Diet. 68, 27–48.
Stryker, W.S., Salvini, S., Stampfer, M.J., Sampson, L., Colditz, G.A., Willett, W.C., 1991. Contributions of specific foods to absolute intake and between-person variation of nutrient consumption. J. Am. Diet. Assoc. 91, 172–178.
Taiwan Department of Health, 2002a. Taiwan Food Composition Database: basic food composition nutrient tables (in Chinese) Retrieved 2004-10-14: http://www.doh.gov.tw/FoodAnalysis/ingredients.htm.
Taiwan Department of Health, 2002b. Taiwan Food Composition Database: principle of classification and coding system (in Chinese) Retrieved 2004-10-14: http://www.doh.gov.tw/FoodAnalysis/account.htm.
Taiwan Department of Health, 2002c. Taiwan Food Composition Database: Vegetable (in Chinese). Retrieved 2004-10-14: http://www.doh.gov.tw/ufile/Doc/蔬菜類_030607.xls.
Vaask, S., Pomerleau, J., Pudule, I., Grinberga, D., Abaravicius, A., Robertson, A., Mckee, M., 2004. Comparison of the Micro-Nutrica Nutritional Analysis program and the Russian Food Composition Database using data from the Baltic Nutrition Surveys. Eur. J. Clin. Nutr. 58, 573–579.
Van Wave, T.W., Decker, M., 2003. Secondary analysis of a marketing research database reveals patterns in dairy product purchases over time. J. Am. Diet. Assoc. 103, 445–453.
Wahlqvist, M.L., Lee, M.S., 2007. Regional food culture and development. Asia. Pac. J. Clin. Nutr. 16, 2–7.
Wikipedia, 2006. Grape. Retrieved 2006-06-21: http://en.wikipedia.org/wiki/Grape.
Zhang, Z.W., Shimbo, S., Miyake, K., Watanabe, T., Nakatsuka, H., Matsuda-Inoguchi, N., Moon, C.S., Higashikawa, K., Ikeda, M., 1999. Estimates of mineral intakes using food composition tables vs measures by inductively-coupled plasma mass spectrometry: Part 1. calcium, phosphorus and iron. Eur. J. Clin. Nutr. 53, 226–232.