

# Adaptation and evaluation of the National Cancer Institute's Diet History Questionnaire and nutrient database for Canadian populations

Ilona Csizmadia<sup>1,\*</sup>, Lisa Kahle<sup>2</sup>, Ruth Ullman<sup>1</sup>, Ursula Dawe<sup>1</sup>, Thea Palmer Zimmerman<sup>3</sup>, Christine M Friedenreich<sup>1</sup>, Heather Bryant<sup>1</sup> and Amy F Subar<sup>4</sup>

<sup>1</sup>Division of Population Health and Information, Alberta Cancer Board, 1331–29 Street NW, Calgary, Alberta, Canada, T2N 4N2; <sup>2</sup>Information Management Services, Inc., Bethesda, MD, USA; <sup>3</sup>Westat, Rockville, MD, USA;

<sup>4</sup>National Cancer Institute, Bethesda, MD, USA

Submitted 26 April 2005; Accepted 25 January 2006

## Abstract

**Background and objective:** Despite assumed similarities in Canadian and US dietary habits, some differences in food availability and nutrient fortification exist. Food-frequency questionnaires designed for the USA may therefore not provide the most accurate estimates of dietary intake in Canadian populations. Hence, we undertook to evaluate and modify the National Cancer Institute's Diet History Questionnaire (DHQ) and nutrient database.

**Methods:** Of the foods queried on the DHQ, those most likely to differ in nutrient composition were identified. Where possible these foods were matched to comparable foods in the Canadian Nutrient File. Nutrient values were examined and modified to reflect the Canadian content of minerals (calcium, iron, zinc) and vitamins (A, C, D, thiamin, riboflavin, niacin, B<sub>6</sub>, folate and B<sub>12</sub>). DHQs completed by 13 181 Alberta Cohort Study participants aged 35–69 years were analysed to estimate nutrient intakes using the original US and modified versions of the DHQ databases. Misclassification of intake for meeting the Dietary Reference Intake (DRI) was determined following analysis with the US nutrient database.

**Results:** Twenty-five per cent of 2411 foods deemed most likely to differ in nutrient profile were subsequently modified for folate, 11% for vitamin D, 10% for calcium and riboflavin, and between 7 and 10% for the remaining nutrients of interest. Misclassification with respect to meeting the DRI varied but was highest for folate (7%) and vitamin A (7%) among men, and for vitamin D (7%) among women over 50 years of age.

**Conclusion:** Errors in nutrient intake estimates owing to differences in food fortification between the USA and Canada can be reduced in Canadian populations by using nutrient databases that reflect Canadian fortification practices.

## Keywords

Food-frequency questionnaire  
Nutrient database  
Fortification  
Misclassification  
Dietary Reference Intake

Food records, 24-hour recalls and food-frequency questionnaires (FFQs) are commonly used to estimate dietary intake in epidemiological research. Each method is characterised by a number of advantages and limitations which need to be considered when choosing a tool for dietary data collection. In large epidemiological studies where measurement of the usual intake of a broad range of nutrients and dietary constituents is the objective, FFQs are usually the most feasible instruments for practical and economic considerations<sup>1</sup>.

The most widely used FFQs are the Block<sup>2</sup>, Willett<sup>3</sup> and more recently the National Cancer Institute (NCI) Diet History Questionnaire (DHQ)<sup>4–6</sup>, which were all developed specifically for US populations. Canadian researchers conducting large dietary studies have also relied on

these FFQs, since few comparable, extensively tested and validated FFQs have been developed specifically for use in Canada.

Some differences are known to exist in food availability and nutrient fortification laws<sup>7,8</sup> between the USA and Canada. Hence Canadian nutritionists and epidemiologists have questioned whether FFQs and nutrient databases designed for US populations can provide valid estimates of nutrient intake in Canadian populations. This concern was addressed in part by Barr *et al.*<sup>9</sup>, who analysed food records with US and Canadian nutrient databases and found overestimates of 5–12% for thiamin, riboflavin, niacin and iron intake with the use of solely US-based food composition data. Comparisons of estimates for pyridoxine, folate, vitamin B<sub>12</sub> and zinc were limited by the

\*Corresponding author: Email ilona.csizmadia@cancerboard.ab.ca

number of foods with missing values in the Canadian nutrient database at that time. More recent comparisons of dietary intake estimates using more complete Canadian data have not been conducted.

The *de novo* development and testing of an FFQ is labour-intensive and costly. In addition, insufficient dietary intake data pertaining to the population for which the FFQ is intended often precludes this approach. In Canada, an FFQ developed by Jain *et al.*<sup>10</sup> in the early 1980s was largely based on dietary intake data collected in observational studies on cancer aetiology. Shatenstein *et al.*<sup>11</sup> recently adapted the Block NCI Health Habits and History Questionnaire using Quebec (a province in Canada) survey data from 1990. At about the same time, we undertook to modify the recently developed and validated US NCI DHQ and its associated nutrient database for use in a newly established longitudinal Canadian cohort, the Alberta Cohort Study (ACS)<sup>12</sup>. The incorporation of cognitive testing methods in the design of the DHQ and the accessibility of the nutrient database for modification were important factors in our decision to use it in the ACS.

The aims of this paper are twofold: first to describe the process of adapting the DHQ and modifying its nutrient database for consistency with Canadian nutrient fortification practices and food availability, and second to describe the impact of using the original US nutrient database versus the modified Canadian database on estimates of dietary intake in the ACS. The proportion of the cohort meeting Dietary Reference Intake (DRI)<sup>13–15</sup> recommendations is used as the criterion to illustrate the potential for misclassification of nutritional intake when food fortification is not accounted for.

## Materials and methods

### DHQ

The DHQ is a cognitive-based FFQ developed by investigators at the US NCI<sup>5,6,16</sup>. It has been described previously and a copy of the original questionnaire and related documentation may be obtained at <http://appliedresearch.cancer.gov/DHQ/index.html>. Briefly, the DHQ queries frequency of intake over the past year for 124 individual food items and asks about portion size for most of these by providing a choice of three ranges. For 44 of the 124 foods, additional embedded questions ask about related factors such as seasonal intake, food type (e.g. low-fat, lean, diet, caffeine-free) and/or fat uses or additions. This results in 255 separate line items on the DHQ.

As the first step in modifying the DHQ for use in Canada, the original DHQ was evaluated for face validity. Each question was examined to ensure that foods appearing on the FFQ were available, consumed in Canada and that important Canadian products were included. Canadian divisions of food companies such as

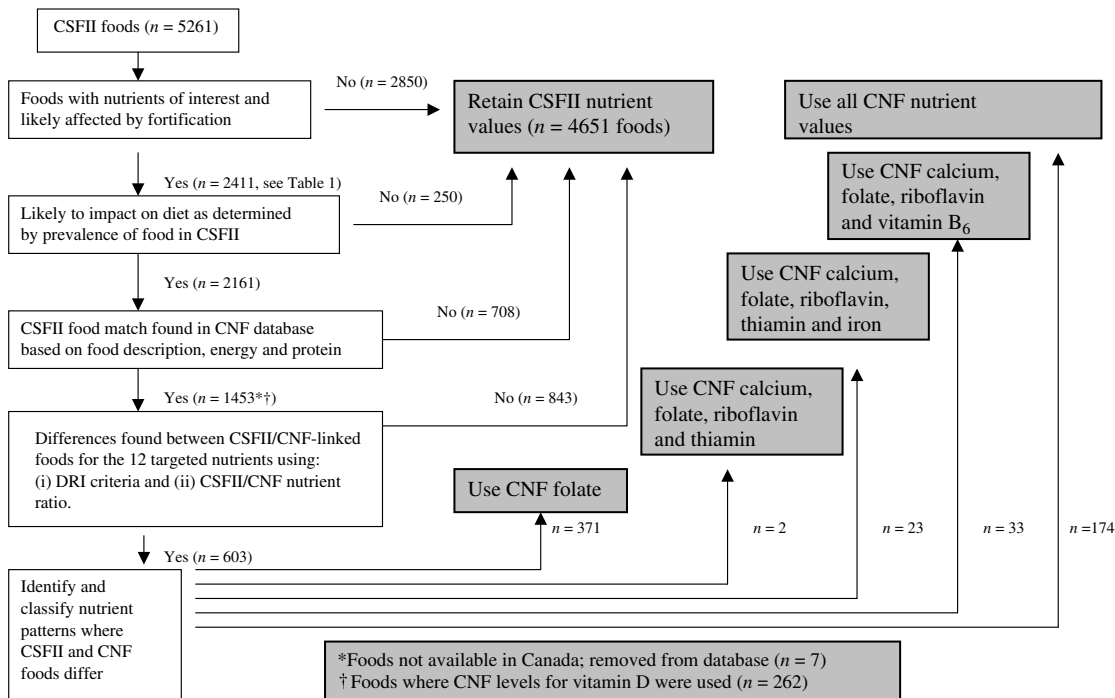
Kellogg's, General Mills Corporation Canada and Minute Maid Canada Inc. were contacted about specific products. In addition, minor modifications were made to the questionnaire to reflect Canadian language usage.

### DHQ and US nutrient database

The US Department of Agriculture's (USDA) 1994–1996 Continuing Survey of Food Intake of Individuals (CSFII) is the primary source of US food and nutrient information for the original DHQ nutrient database<sup>17</sup>. The USDA Nutrient Database for Standard Reference (SR11) was the food composition source for CSFII 1994–96. Using 5261 unique foods that were reported in the CSFII by adults 19 years of age or older, food groups were created to represent foods of similar usage and nutritional value. Details regarding methods by which these CSFII food groups were used to create the nutrient database for the DHQ are detailed elsewhere<sup>6</sup>. Briefly, however, for each food group, estimates of nutrient composition were generated, by gender and portion size, by calculating mean nutrient estimates with foods weighted by frequency of consumption as reported in the CSFII. More recently, carotenoids, tocopherols, vitamin D and other important nutrients have been added to the database using data from the Nutrition Data System for Research<sup>18</sup>.

An overview of the steps involved in the evaluation and modification of the CSFII-based DHQ nutrient database is presented in Fig. 1 and described below. Several criteria were employed in determining which food groups were most likely to differ in nutrient content due to differences in fortification practices between the USA and Canada. First, food groups consisting mainly of unprocessed foods such as fruits and vegetables were not considered further, whereas groups consisting mostly of foods that had undergone a manufacturing process (e.g. breads, cereals and baked products) were retained for further examination. For the second criterion, food groups considered important sources of selected vitamins and minerals were retained. The following vitamins and minerals were identified *a priori* as those most likely to be affected by fortification<sup>7</sup> and manufacturing practices and also most relevant in studies assessing the associations between diet and chronic disease: vitamins A, C, B<sub>6</sub>, B<sub>12</sub> and D, thiamin, riboflavin, niacin, folate, calcium, iron and zinc<sup>13–15,19</sup>. Following this process the food groups retained for further examination comprised 2411 individual foods.

The final criterion involved identifying which of the 2411 foods were most likely to have an important impact on overall weighted mean nutrient estimates. This involved identifying foods by unique food codes for which reported frequencies of intake in the CSFII were less than 1% of the total frequency within each group. These foods were then excluded, leaving 2161 foods of interest.



**Fig. 1** Steps involved in the evaluation, modification and linking of foods in the CSFII-based DHQ nutrient database to foods in the CNF database, to account for nutrient differences due to US and Canadian food fortification. CSFII – 1994–1996 Continuing Survey of Food Intake of Individuals; DHQ – National Cancer Institute’s Diet History Questionnaire; CNF – Canadian Nutrient File; DRI – Dietary Reference Intake

### Canadian Nutrient File (CNF)

Food composition data for Canada were obtained from the Canadian Nutrient File (CNF, Version 2001b)<sup>20</sup>. This large relational database, compiled by the Nutrition Research Division at Health Canada, is the most comprehensive and valid source of Canadian food composition information. The database contains average values for 112 nutrients in 4943 foods available in Canada. Although 68% of the foods in the CNF are derived unchanged from the USDA Nutrient Database for Standard Reference (SR13), the remaining are foods with nutrient values modified to reflect fortification and regulatory standards specific to Canada or are brand names of foods available only in Canada.

### Linking CSFII foods to CNF database foods

Comparison of CSFII foods with CNF foods necessitated the linkage of CSFII foods with corresponding CNF database foods. Because an automated system was not available we matched each of the 2161 CSFII foods with directly comparable foods in the CNF database on a food-by-food basis. This matching was achieved by comparing food descriptions and energy and protein content per 100 g of food. A match was considered acceptable if, in addition to the descriptive match, the ratios of CSFII to CNF protein and energy per 100 g of food were between 0.8 and 1.2. This criterion, previously used to guide the matching of foods in the CSFII database to foods in the Nutrition Data System for Research for the purposes of adding carotenoids to the DHQ nutrient database<sup>18</sup>, was

based on informed nutritional judgement that it was able to discern poor matches.

When a close match was not possible, we sought to match the CSFII food with the closest reasonable food from the CNF. A numerical system was used to code the quality of the match based on description alone. This match code system denotes conditions where the energy and protein values are within the range described above but where the food description indicates that the match may be questionable on other characteristics of that food, as described in Table 1. These codes were then used to sort the linked foods by quality of match during the next steps that involved the detailed examination of linked foods for differences in the 12 nutrients of interest with respect to fortification.

### Evaluation of CSFII/CNF-linked foods for nutrient differences

The first step in the examination of the CSFII/CNF-linked foods was to ensure that comparisons were made between nutrients in foods where the nutrient amounts were high enough to potentially impact on daily dietary intake. If at least one of the two linked foods in a match contained at least 10% of the DRI for a specific nutrient (5% for folate and calcium) per 100 g of food, the nutrient was coded as eligible for comparison. The criterion of CSFII/CNF ratios 0.8 to 1.2 was then used to judge whether or not there were important differences per 100 g portions of food for each of the 12 nutrients mentioned above. Nutrients were

**Table 1** Frequency and descriptions of match quality for CSFII/CNF database-linked foods matched on energy and protein content and individual food descriptions

Description of match	Foods	
	<i>n</i>	%
Acceptable match	845	35
No suitable match	709	29
Match questionable for reasons other than those described below	413	17
Food frequency in CSFII < 1% remained unmatched	250	10
CNF food missing one descriptive component	66	3
Fat description does not match well	57	2
Sodium description does not match well	41	2
Breads and rolls toasted in CSFII but not in CNF	16	< 1
Calcium description does not match well	2	< 1
Sweetener description does not match well	5	< 1
CSFII food missing one descriptive component	3	< 1
Juice with added vitamin C in CSFII but not in CNF	4	< 1
Total	2411	100

CSFII – 1994–1996 Continuing Survey of Food Intake of Individuals; CNF – Canadian Nutrient File.

classified and coded to indicate membership in one of the following four categories of CSFII/CNF ratio: (1) 0.8 to 1.2 inclusive, as an indicator of negligible nutrient differences; (2) less than 0.8 (nutrient level was higher in CNF food); (3) greater than 1.2 (nutrient level was higher in CSFII food); and (4) when CNF nutrient value was missing. The prevalence of CSFII/CNF nutrient ratio codes across food categories was then used as a guide to identify nutrient pattern differences between CSFII/CNF-linked foods. Common patterns of nutrient differences were identified and grouped according to the required modifications. Where nutrient ratios were less than 0.8 or greater than 1.2, the next step was to determine if differences were due to fortification or to discrepancies in ingredients. Based on an examination of food descriptions again, decisions were made whether or not to substitute CNF nutrient values for CSFII values. Since fortification is nutrient-specific, decisions for substitution were also nutrient-specific such that a food in the modified Canadian database could in the end have nutrients drawn from both the CSFII and CNF databases. For example, if a food differed only in folate level, the CSFII database could be the source of nutrients for all of the nutrients except for folate, for which the CNF would be the source.

Where nutrient values for a food were missing in the CNF database, the CSFII values were retained. Exceptions to this rule were the breakfast cereal groups. If the CSFII value exceeded the maximum allowable level of fortification specified in the Canada Food and Drug Act (CFDA)<sup>7</sup>, the maximum CFDA nutrient value was imputed. The maximum allowable levels of nutrients that can be added to 100 g of cereal are as follows: vitamin B<sub>6</sub>, 0.6 mg; thiamin, 2.0 mg; niacin, 4.8 mg; iron, 13.3 mg; folate, 60.0 µg; magnesium, 160 mg; zinc, 3.5 mg. The CFDA does not specify maximum amounts of calcium and riboflavin that can be added to cereals. Cereals in Canada are not

fortified with vitamins C, B<sub>12</sub>, A and D<sup>7</sup>, nor do they generally occur naturally in cereal. Hence, when these nutrient values were missing for cereals in the CNF database, CSFII values were replaced with '0'.

Once decisions for nutrient modifications were finalised, the Canadian DHQ database was generated. For the creation of this database, the CSFII data for portion size and frequency of food intake were employed with methods similar to those used to create the original US DHQ database<sup>6</sup>.

### **Comparison of estimates of nutrient intake using the US and Canadian nutrient databases**

The ACS is a large prospective cohort, with several waves of ongoing recruitment. Men and women between 35 and 69 years of age, residing in the province of Alberta, who have not been diagnosed with cancer, are invited to participate. At baseline all participants are requested to complete the Canadian version of the DHQ. The ethics committees of the Alberta Cancer Board and the University of Calgary granted ethical approval for the ACS.

At the time of this study, 13 181 completed questionnaires were available from participants recruited between 2001 and 2003. Nutritional supplements were also ascertained using the DHQ; however, for this analysis we estimated dietary intakes excluding supplements since we were interested in differences in estimates of nutritional intake using two country-specific nutrient databases.

Questionnaires were excluded if the reported total daily energy intake was judged to be biologically implausible (<800 or >4200 kcal for men; <600 or >3500 kcal for women) as proposed by other authors<sup>4,21</sup>. This rule resulted in 543 exclusions (304 men and 239 women), leaving 12 638 questionnaires available for analysis.

### **Statistical analyses**

Nutrient analyses of the DHQs were carried out using Diet\*Calc (Version 1.4.2) software. The Diet\*Calc data dictionary was modified for the Canadian version of the DHQ. Estimates of nutrient intakes were determined using the original CSFII database and again using the CNF-modified nutrient database. Means, standard deviations (SDs), medians and the proportions of individuals meeting the age-specific DRI for the 12 relevant nutrients<sup>14,13,22</sup> were calculated for results from the two databases.

Nutrient mean differences of at least 5% of the Canadian estimated mean intakes were tested statistically using a *t*-test for paired data to determine whether or not the mean differences were significantly different from zero. Continuous nutrient variables were transformed to a logarithmic scale in order to normalise distributions for the statistical testing of mean differences. McNemar's test was used to ascertain significant differences between proportions of individuals classified as meeting the DRI when intakes were estimated using the two nutrient databases.

Testing was carried out for nutrients where estimates of proportions differed by 5% or more. Since DRIs are sex- and age-specific, analyses were stratified by these variables as indicated<sup>13–15</sup>. All analyses were carried out using Statistical Analysis System (SAS) version 8.2 (SAS Institute Inc., Cary, NC, USA).

## Results

### *Modification of the DHQ*

Two line items were removed from the original DHQ as they referred to products currently not available in Canada: highly fortified cereals (Total, Product 19 and Right Start) and potato chips made with Olean® or Olestra® fat substitutes.

A number of items not on the original version of the DHQ were added in order to more completely capture foods and nutrients of interest for a Canadian context. An embedded separate question that queried intake of orange and grapefruit juice fortified with calcium was added to the question pertaining to orange and grapefruit juice intake. A separate question was added to ascertain the intake of supplemental vitamin D consumed with or without calcium supplements but in addition to the amount taken in multivitamin preparations. Splenda® (sucralose), a modified sugar that is not fully absorbed or metabolised and is 600 times sweeter than sucrose<sup>23</sup>, was added as a third possible choice in the category of artificial sweeteners.

A number of language modifications were made to reflect Canadian language usage; e.g. 'grits' were replaced with 'cream of wheat', 'biscuits' with 'baking powder biscuits', 'catsup' was spelled 'ketchup' and 'fatback' was deleted. Eleven such minor language modifications were made.

### *CSFII/CNF-linked foods*

The algorithm for matching CSFII database foods to CNF database foods is presented in Fig. 1 and the descriptions of matching quality are listed in Table 1. Of the 2161 CSFII database foods for which matching was sought, matches were not found for 708 foods. Of the 1453 matched foods, 845 were judged to be acceptable. The remaining 608 were questionable due to discrepancies in food descriptions for type of sweetener or fat, vitamin C, sodium or calcium content and 'other' discrepancies (Table 1). This latter group includes 413 potential mismatches associated with different methods of preparation, e.g. homemade versus commercially prepared foods or composition differences as indicated by the ingredients listed in the food descriptions. For example, 'chicken broth, with tomato, home recipe' in the CSFII database is matched with 'soup, broth, chicken, canned, condensed, and water' in the CNF database.

### *Modification of the DHQ nutrient database*

The 1453 CSFII/CNF linked foods were examined using the match descriptions in Table 1, DRI criteria for determining relevant nutrient densities in a food and CSFII/CNF nutrient ratios as described above. Six hundred and three linked foods were identified as requiring one or more nutrient substitutions from the CNF and were then grouped according to the modifications required. The type of modification and the number of foods within each group are outlined in Fig. 1.

Twenty-five per cent of the 2411 foods originally identified as potentially requiring modification were in the end modified to reflect Canadian levels of nutrient content. All of the 25% were modified for folate and 7–10% were modified for riboflavin, calcium, iron, zinc, niacin, thiamin and vitamins A, B<sub>6</sub>, B<sub>12</sub> and C. In addition, 11% were also modified for vitamin D.

Foods most frequently requiring nutrient modifications were ready-to-eat cereals and commercially prepared baked products. In addition, milk replacements (rice and soy) were fortified with vitamins D, A, B<sub>12</sub>, riboflavin, zinc and calcium in Canada but not in the CSFII 1994–96 nutrient database. The fortification of margarine with vitamin D in Canada but not in the USA is also noteworthy.

### *Comparison of dietary intake estimates using the CSFII nutrient database and the CNF-modified database*

DHQs completed by 4979 men and 7659 women, aged 35–69 years, living in Alberta, Canada, were analysed and compared for differences in nutrient intake using the original CSFII-based nutrient database and the CNF-modified database for the DHQ. Means, SDs and medians for reported daily intakes of energy, protein, carbohydrate, dietary fibre, fat and the 12 vitamins and minerals modified in the database are presented for men and women in Table 2. Since macronutrients were not modified from original CSFII values, similar values were obtained for estimates of energy, protein, fat, dietary fibre and carbohydrate with both databases. The addition of calcium-fortified orange juice as a separate line item to the Canadian version of the DHQ explains the slight difference in the estimates of carbohydrate and energy. For all nutrients and energy, median values were below mean values reflecting the usual lack of normality in the distribution of dietary intakes.

For men, the largest mean differences in nutrient intakes using the two databases were for folate and vitamin A. Compared with the CNF-modified database, CSFII database mean estimates for these nutrients were higher by 27 µg (7%) and 111 retinol equivalents (RE) (7%), respectively (Table 2). Assuming that the CNF-modified database is a better estimate of the nutrient content of Canadian foods, at the individual level the CSFII-based database yields intakes for folate that can range from an overestimate of 447 µg to an underestimate of 72 µg. The

**Table 2** Means (SD) and medians for nutrient intakes of men and women using CSFII-based and CNF-based nutrient databases to analyse DHQs

Nutrient	Men				Women				
	CNF		CSFII		CNF		CSFII		
	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median	
<i>Ages 35 to 69 years</i>		<i>n = 4979</i>				<i>n = 7659</i>			
Energy (kcal)	2126 (720)	2042	2123 (720)	2038	1617 (549)	1536	1612 (547)	1532	
Protein (g)	82 (32)	77	82 (32)	77	64 (24)	61	64 (24)	61	
Total fat (g)	78 (33)	73	78 (33)	73	58 (25)	54	58 (25)	54	
Dietary fibre (g)	20 (9)	18	20 (9)	18	18 (8)	17	18 (8)	17	
Carbohydrate (g)	262 (94)	248	261 (94)	247	211 (77)	200	209 (77)	198	
Folate (µg)	357 (144)	335	384 (153)*	361	311 (129)	290	331 (136)*	311	
Niacin (mg)	24 (9)	23	25 (9)	24	18 (7)	17	19 (7)	18	
Riboflavin (mg)	2.0 (0.8)	1.8	2.0 (0.8)	1.9	1.6 (0.7)	1.5	1.7 (0.7)	1.6	
Thiamin (mg)	1.7 (0.6)	1.6	1.7 (0.6)	1.6	1.3 (0.5)	1.3	1.3 (0.5)	1.3	
Vitamin A (RE)	1338 (814)	1125	1449 (839)*	1243	1383 (908)	1149	1454 (920)*	1222	
Vitamin B <sub>12</sub> (µg)	5.5 (3.0)	4.8	5.7 (3.0)	5.1	4.0 (2.6)	3.5	4.1 (2.7)	3.7	
Vitamin C (mg)	144 (97)	121	148 (98)	126	145 (91)	126	147 (91)	128	
Zinc (mg)	12.8 (5.8)	11.6	13.2 (5.8)	11.9	9.5 (4.1)	8.7	9.8 (4.2)	9.0	
<i>Ages 35 to 50 years</i>		<i>n = 2606</i>				<i>n = 4107</i>			
Iron (mg)	15.5 (5.4)	14.8	16.4 (6.0)*	15.5	12.0 (4.3)	11.5	12.6 (4.7)*	11.9	
Calcium (mg)	989 (501)	883	988 (502)	880	830 (420)	746	826 (419)	741	
Vitamin D (µg)	5.4 (3.5)	4.5	5.3 (3.5)	4.4	4.3 (2.8)	3.5	4.1 (2.8)	3.4	
Vitamin B <sub>6</sub> (mg)	2.1 (0.8)	2.0	2.2 (0.9)*	2.1	1.7 (0.7)	1.6	1.8 (0.7)*	1.7	
<i>Ages 51 to 69 years</i>		<i>n = 2373</i>				<i>n = 3552</i>			
Iron (mg)	14.5 (5.2)	13.8	15.3 (5.7)*	14.6	11.7 (4.2)	11.1	12.3 (4.6)*	11.6	
Calcium (mg)	837 (430)	743	840 (430)	744	765 (398)	680	762 (395)	675	
Vitamin D (µg)	4.7 (3.1)	3.9	4.5 (3.1)	3.8	4.1 (2.7)	3.4	3.8 (2.6)*	3.1	
Vitamin B <sub>6</sub> (mg)	2.0 (0.8)	1.9	2.1 (0.8)*	2.0	1.7 (0.6)	1.6	1.8 (0.7)*	1.7	

SD – standard deviation; CSFII – 1994–1996 Continuing Survey of Food Intake of Individuals; CNF – Canadian Nutrient File; DHQ – National Cancer Institute's Diet History Questionnaire; RE – retinol equivalents.

\* CSFII-based database mean nutrient intake at least 5% more or less than the estimated mean intake for the nutrient using the CNF-based Canadian database. Mean differences for these nutrients were significantly different from zero ( $P < 0.0001$ ).

estimates for vitamin A can range from an overestimate of 1658 RE to an underestimate of 600 RE. This illustrates the extent to which food selection at the individual level can impact on the magnitude of measurement error that may occur with the use of a database that does not correctly reflect Canadian food fortification.

In women, the greatest difference in mean intake occurred for vitamin D among women 51 years of age and older. Compared with the CNF-based database, mean vitamin D intake was estimated to be 0.3 µg (7%) lower with the CSFII-based database. At the individual level, differences ranged from an overestimate of 1.7 µg to an underestimate of 6.4 µg of vitamin D.

All differences in mean estimates of nutrient intakes that were greater than 5% of the Canadian estimate (folate, vitamin A, iron and vitamin B<sub>6</sub> for both men and women, and vitamin D among women 51 years of age and older) were significantly different from zero ( $P < 0.0001$ ).

The proportions of men and women meeting DRI recommendations for the 12 nutrients affected by fortification using the original CSFII and CNF-modified nutrient databases are reported in Table 3. With the CNF-modified database, only 32% of men met the folate DRI of 400 µg; however, with the original CSFII-based database, 39% (366 additional men) were classified as having achieved this level of intake. For vitamin A intake the

magnitude of misclassification was similar, with 7% more men classified as meeting or exceeding the DRI for vitamin A when the original CSFII-based database was used. Six per cent of men and women aged 51–70 years were misclassified with respect to meeting the DRI for vitamin B<sub>6</sub>. Differences in proportion estimates with respect to meeting the DRIs were statistically significant ( $P < 0.0001$ ) for nutrients with misclassification of 5% or more.

### Discussion

In the present paper we describe modifications made to the NCI's DHQ and nutrient database in order to improve estimates of dietary intake in Canadian populations. We undertook extensive examination of foods in the original DHQ nutrient database, to identify differences in food composition as a result of different food fortification policies in the USA and Canada. We also demonstrate the potential for misclassifying nutrient intake in a Canadian population when a nutrient database that does not account for Canadian food fortification practices is used.

Vitamins and minerals have been added to foods in the USA and Canada since the availability of synthetic vitamins in the 1940s<sup>24</sup>. Generally, the two countries have had similar objectives for programmes and policies of food fortification. These include adding nutrients to foods for

**Table 3** Percentage of men and women with intake meeting the DRI for nutrients using CSFII-based and CNF-based nutrient databases to analyse DHQs

Nutrient	Men ( <i>n</i> = 4979)			Women ( <i>n</i> = 7659)		
	DRI	Percentage with intake meeting DRI		DRI	Percentage with intake meeting DRI	
		CNF	CSFII		CNF	CSFII
<i>Ages 31 to 70 years</i>						
Folate ( $\mu\text{g}$ )	400	32	39*	400	20	25*
Niacin (mg)	16	82	86	14	70	74
Riboflavin (mg)	1.3	80	83	1.1	76	79
Thiamin (mg)	1.2	76	76	1.1	65	64
Vitamin A (RE)	900	66	73*	700	80	84
Vitamin B <sub>12</sub> ( $\mu\text{g}$ )	2.4	92	93	2.4	77	80
Vitamin C (mg)	90	69	71	75	81	82
Zinc (mg)	11	55	58	8	59	62
<i>Ages 31 to 50 years</i>						
Iron (mg)	8	95	95	18	9	13
Calcium (mg)	1000	40	40	1000	28	28
Vitamin D ( $\mu\text{g}$ )	5	44	42	5	30	28
Vitamin B <sub>6</sub> (mg)	1.3	87	89	1.3	69	73
<i>Ages 51 to 70 years</i>						
Iron (mg)	8	92	93	8	82	83
Calcium (mg)	1200	16	17	1200	13	13
Vitamin D ( $\mu\text{g}$ )	10	5.4	4.9	10	3.1	2.8
Vitamin B <sub>6</sub> (mg)	1.7	59	65*	1.5	56	62*

DRI – Dietary Reference Intake; CSFII – 1994–1996 Continuing Survey of Food Intake of Individuals; CNF – Canadian Nutrient File; DHQ – National Cancer Institute's Diet History Questionnaire; RE – retinol equivalents.

\* Differences of at least 5% in the estimates of proportions of individuals classified as meeting the DRI for a given nutrient using the CNF-based and CSFII-based databases. These proportions were significantly different from each other ( $P < 0.0001$ ).

purposes of preventing or correcting nutrient deficiencies and restoring nutrients that are lost during manufacturing and processing. While both countries have also targeted similar nutrients for fortification and enrichment, there are some differences with respect to the specific foods and permissible levels to which nutrients can be added. For example, Canadian regulations allow for the addition of vitamin D in specified amounts to all types of milk and milk substitutes (rice and soy) and to margarine, but not to cereals. In Canada, the principle of nutritional equivalence has determined the permitted level of plant-based beverage nutrient fortification, since these products are considered to be interchangeable with fluid milk<sup>24</sup>. In the USA vitamin D is added to cow's milk, ready-to-eat cereals and more recently to milk substitutes, but not to margarine.

The restoration of B vitamins and the addition of folic acid and iron to flour and refined cereal products are under strict regulatory control in Canada<sup>7</sup>, where maximum levels are specified. In the case of folic acid, a maximum of 60  $\mu\text{g}/100\text{g}$  may be added to ready-to-eat cereals in Canada, while in the USA a minimum level of 140  $\mu\text{g}/100\text{g}$  is specified but maximum levels are not<sup>8</sup>. Hence amounts greatly

exceeding this level have been reported for US ready-to-eat cereals<sup>25</sup>. In our comparisons of identical ready-to-eat cereal products available in both the USA and Canada, levels of folate differed by as much as five times in some cases. In addition, vitamins A, C and D may be added to ready-to-eat cereals in the USA but not in Canada.

At the population level, our results appear to indicate minor differences in estimates of mean nutrient intakes with the use of the CNF-based database versus the original CSFII-based nutrient database. The full impact of these differences on nutrition study results, however, may be subtle and a challenge to evaluate. For example, in a subgroup of men in the ACS reporting a habitual intake of cereal at least three times per week, 18% were classified (data not shown) as having a folate intake of at least 400  $\mu\text{g}$  (the DRI for folate) compared with only 7% among all men when the CSFII-based database was used. The extent to which this magnitude of misclassification could impact on measures of effect in aetiological studies of diet and disease in subgroups of the population needs careful consideration. We therefore suggest that at the individual level and among certain subgroups with particular dietary habits, there is great potential to reduce misclassification of nutrient intake and dietary measurement error with the use of a database that more closely reflects food fortification practices.

The presence of systematic and random measurement error in dietary studies is well recognised<sup>26–28</sup>. The development of methodological approaches to quantify and minimise errors that result from the incomplete reporting of foods and misreporting of frequency of intake and portion size continues to be a challenge to nutritional epidemiologists. Striving to achieve accuracy in the food composition data of nutrient databases is an effective strategy in minimising a quantifiable source of error. We have adapted the NCI's DHQ nutrient database to more accurately reflect the nutrient composition of foods affected by fortification in Canada. Estimates of dietary intake using this CNF-modified nutrient database will potentially reduce the misclassification in estimates of intake for important nutrients relevant to epidemiological research in Canada.

A few limitations of this work need to be acknowledged. About one-third of the foods that were candidates for matching with foods in the CNF remained unmatched because comparable foods were not found in the CNF database. These foods are largely desserts, baked products and mixed meals, some of which contain fortified ingredients. Although this incomplete matching of foods is of concern, we believe that the most commonly available and most frequently consumed foods in Canada have been accounted for.

Another limitation of our work is the reliance on a US dietary survey for our information pertaining to portion size and prevalence of food consumption. A large cross-sectional survey of 35 000 Canadians (Canadian

Community Health Survey, Cycle 2.2) has recently been completed by Statistics Canada. Within the next year 24-hour dietary recall data collected from participants in this survey will be available for analysis. We anticipate that these data will provide us with the opportunity to examine our assumptions about the similarities between the USA and Canada with regard to various patterns of food intake, e.g. food selection, frequency of food intake and portion size.

At present, however, our efforts provide Canadian nutritionists and epidemiologists with a well-tested FFQ that was designed using cognitive-based methods and food consumption data from a large free-living population. The detailed description of the extensively modified nutrient database and the comparisons of dietary intake estimates using the original CSFII and CNF-modified nutrient databases will better equip investigators with the information they need when deciding on a method of dietary data collection and analysis.

### Acknowledgements

**Funding:** I.C. was supported by Alberta Cancer Foundation and Canadian Institutes of Health Research Fellowships. C.M.F. was supported by a CIHR New Investigator Award and an Alberta Heritage Foundation for Medical Research Health Scholar Award. Funding for the Alberta Cohort Study was provided by the Alberta Cancer Board's New Initiatives Program and the Alberta Cancer Foundation.

**Contributors:** I.C., C.M.F. and A.F.S. developed the design and analysis of the study and interpreted the data. I.C., U.D., C.M.F., A.F.S., H.B. and R.U. were responsible for the acquisition of data. I.C. drafted the manuscript, and C.M.F. and A.F.S. contributed to important intellectual content and critical revisions. Technical expertise and programming support were provided by I.C., L.K. and T.P.Z. Funding for the study was obtained by C.M.F. and H.B.

**Conflicts of interest:** None of the authors have any conflict of interest to declare.

**Thanks:** The authors would like to thank Ms Josephine Deeks of the Nutrition Research Division of Health Canada for her assistance in providing background information pertaining to the Canadian Nutrient File and Dr Paula Robson of the Alberta Cancer Board for her helpful comments in the preparation and editing of the manuscript. Ms Deeks and Dr Robson have provided permission to be acknowledged for their contribution.

### References

- 1 Nelson M, Bingham SA. Assessment of food composition and nutrient intake. In: Margetts BM, Nelson M, eds. *Design Concepts in Nutritional Epidemiology*. Oxford: Oxford University Press, 1997.
- 2 Block G, Hartman AM, Dresser CM, Carroll MD, Gannon J, Gardner L. A data-based approach to diet questionnaire design and testing. *American Journal of Epidemiology* 1986; **124**: 453–69.
- 3 Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, *et al.* Reproducibility and validity of a semiquantitative food frequency questionnaire. *American Journal of Epidemiology* 1985; **122**: 51–65.
- 4 Subar AF, Thompson FE, Kipnis V, Midthune D, Hurwitz P, McNutt S, *et al.* Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. *American Journal of Epidemiology* 2001; **154**: 1089–99.
- 5 Thompson FE, Subar AF, Brown CC, Smith AF, Sharbaugh CO, Jobe JB, *et al.* Cognitive research enhances accuracy of food frequency questionnaire reports: results of an experimental validation study. *Journal of the American Dietetic Association* 2002; **102**: 212–25.
- 6 Subar AF, Midthune D, Kulldorff M, Brown CC, Thompson FE, Kipnis V, *et al.* Evaluation of alternative approaches to assign nutrient values to food groups in food frequency questionnaires. *American Journal of Epidemiology* 2000; **152**: 279–86.
- 7 Health Canada. *Canadian Food and Drugs Act and the Foods and Drug Regulations. Part B* [online], 1998. Available at <http://laws.justice.gc.ca/en/F-27/C.R.C.-c.870/123128.html>. Accessed 22 December 2005.
- 8 Food and Drug Administration. Food standards: amendment of the standards of identity for enriched grain products to require addition of folic acid. *Federal Register* 1996; **61**: 8781–897.
- 9 Barr SI, Kwan S, Janelle KC. Nutrient analysis using computer programs: comparison of a Canadian and an American database. *Journal of the Canadian Dietetic Association* 1994; **55**: 29–32.
- 10 Jain MG, Harrison L, Howe GR, Miller AB. Evaluation of a self-administered dietary questionnaire for use in a cohort study. *American Journal of Clinical Nutrition* 1982; **36**: 931–5.
- 11 Shatenstein B, Nadon S, Godin C, Ferland G. Development and validation of a food frequency questionnaire. *Canadian Journal of Dietetic Practice and Research* 2005; **66**: 67–75.
- 12 Bryant H, Robson PJ, Ullman R, Friedenreich C, Dawe U. Population-based cohort development in Alberta, Canada: a feasibility study. *Chronic Diseases in Canada* 2006; **27**: 55–63.
- 13 Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. *Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D and Fluoride*. Washington, DC: National Academy Press, 2000.
- 14 Panel on Dietary Antioxidants and Related Compounds. *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids*. Washington, DC: National Academy Press, 2000.
- 15 Standing Committee on the Scientific Evaluation of Dietary Reference Intakes. *Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B<sub>6</sub>, Folate, Vitamin B<sub>12</sub>, Pantothenic Acid, Biotin, and Choline*. Washington, DC: National Academy Press, 1998.
- 16 Subar AF, Thompson FE, Smith AF, Jobe JB, Ziegler RG, Potischman N, *et al.* Improving food frequency questionnaires: a qualitative approach using cognitive interviewing. *Journal of the American Dietetic Association* 1995; **95**: 781–8.
- 17 Tippet KS, Cypel YS. *Design and Operation: The Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey, 1994–96*. NFS Report No. 96-1. Washington, DC: US Department of Agriculture, 1998.
- 18 Dixon L, Zimmerman TP, Kahle LL, Subar AF. Adding carotenoids to the NCI Diet History Questionnaire Database. *Journal of Food Composition and Analysis* 2003; **16**: 269–80.



- 19 World Cancer Research Fund (WCRF)/American Institute of Cancer Research (AICR). *Diet, Nutrition, and the Prevention of Cancer: A Global Perspective*. Washington, DC: WCRF/AICR, 1997.
- 20 Nutrition Research Division, Health Canada. *Canadian Nutrient File, 2001b* [online], 2002. Available at [http://www.hc-sc.ca/food-aliment/ns-sc/nr-rn/surveillance/cnf-fcen/e\\_index.html](http://www.hc-sc.ca/food-aliment/ns-sc/nr-rn/surveillance/cnf-fcen/e_index.html). Accessed 9 December 2004.
- 21 Hung HC, Joshipura KJ, Jiang R, Hu FB, Hunter D, Smith-Warner SA, et al. Fruit and vegetable intake and risk of major chronic disease. *Journal of the National Cancer Institute* 2004; **96**: 1577–84.
- 22 Panel on Micronutrients. *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium and Zinc*. Washington, DC: National Academy Press, 2002.
- 23 Binns NM. Sucralose – all sweetness and light. *Nutrition Bulletin* 2003; **28**: 53–8.
- 24 Interdepartmental Government Working Group. *Addition of Vitamins and Minerals to Foods: Proposed Policy Recommendations*. Ottawa: Bureau of Nutritional Sciences, Food Directorate, Health Protection Branch, Health Canada, 1999.
- 25 Whittaker P, Tufaro PR, Rader JI. Iron and folate in fortified cereals. *Journal of the American College of Nutrition* 2001; **20**: 247–54.
- 26 Subar AF, Kipnis V, Troiano RP, Midthune D, Schoeller DA, Bingham S, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *American Journal of Epidemiology* 2003; **158**: 1–13.
- 27 Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, et al. Structure of dietary measurement error: results of the OPEN biomarker study. *American Journal of Epidemiology* 2003; **158**: 14–21.
- 28 Fraser GE. A search for truth in dietary epidemiology. *American Journal of Clinical Nutrition* 2003; **78**: 521S–5S.