Validity and reliability of the Block98 food-frequency questionnaire in a sample of Canadian women

Beatrice Boucher1,*, Michelle Cotterchio1,2, Nancy Kreiger1,2,3, Victoria Nadalin1, Torin Block4 and Gladys Block5

1Division of Preventive Oncology, Cancer Care Ontario, 620 University Avenue, Toronto, Ontario, Canada, M5G 2L7; 2Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada; 3Department of Nutritional Sciences, University of Toronto, Toronto, Ontario, Canada; 4Block Dietary Data Systems, Berkeley, CA, USA; 5Department of Public Health, University of California, Berkeley, CA, USA

Submitted 14 December 2004: Accepted 18 May 2005

Abstract

Objective: To assess the validity and reliability of the most recent adaptation of Block’s full-diet food-frequency questionnaire (FFQ) among a sample of Canadian women.

Design: Participants completed a self-administered FFQ (FFQ1), two unannounced 24-hour recalls (weekday and weekend) and a second FFQ (FFQ2) between October 2003 and February 2004. FFQs and recalls were analysed for 32 nutrients using Block Dietary Data Systems and the University of Minnesota’s Nutrient Data System. Mean and median intakes were computed, along with crude and deattenuated Pearson correlation coefficients between FFQ1 and the average of two recalls (validity) and between FFQ1 and FFQ2 (reliability).

Setting: Ontario, Canada.

Subjects: A random population-based sample (n = 166) of women aged 25 to 74 years.

Results: One hundred and fifteen (69%) women completed FFQ1, 96 completed FFQ1 and both recalls, and 93 completed both FFQs, about 56 days apart. Mean intakes were similar for most nutrients. FFQ reliability was high, with Pearson correlation coefficients having a median of 0.75, ranging from 0.57 to 0.90 (macronutrients) and from 0.65 to 0.88 (micronutrients from supplements and food). FFQ validity was moderate to high, with deattenuated Pearson correlation coefficients having a median of 0.59, ranging from 0.11 to 0.73 (macronutrients) and from 0.50 to 0.76 (micronutrients from supplements and food). Our micronutrient correlations were similar to or higher than those of other studies that included supplements. Two correlations, 0.40 were associated with fats.

Conclusions: The validity and reliability of this full-diet version of the Block FFQ were moderate to high, supporting its use in future studies among Canadian women.

Food-frequency questionnaires (FFQs) provide the most practical and economical method for collecting data on ‘usual’ dietary intake in population-based epidemiological studies. In aetiological research, it is preferable to use FFQs that query the whole diet rather than a limited number of foods for specific hypotheses, since comprehensiveness will improve the ability to adjust for confounding nutrients and energy intake, when appropriate, and to explore future dietary hypotheses.

Many comprehensive FFQs have been developed for epidemiological research and, among these, the Block FFQ is one of the most widely used. Originally developed at the US National Cancer Institute in 1984, it has undergone a number of revisions to reflect changes in consumption and improvements in FFQ design. Although a new FFQ is in development (Torin Block, personal communication), the most recent full-diet version at the time of this study was developed in 1998 to incorporate dietary and questionnaire changes suggested by American national consumption data (from the Third National Health and Nutrition Examination Survey (NHANES III)) and cognitive research.

Different versions of the Block FFQ have been validated by Block or others to allow comparison with other FFQs or to assess FFQ performance for specific research hypotheses or groups. This approach follows guidelines to validate an FFQ after its development, or whenever it has been substantially changed or used with a new population. Validation involves assessing how well an FFQ measures what it was designed to measure and requires comparison with a ‘more accurate’ reference measure of intake, usually multiple
dietary recalls, records or biomarkers. Validation correlations vary with the nutrient, but typically range from 0.40 to 0.70.\textsuperscript{3,4,22,23}

The current study is the first to report on the validation of the most recent full-diet Block FFQ in common use, Block98, and compares intakes from the FFQ with short-term dietary recall data in a North American population. Additionally, this is the first time the Block FFQ has been validated in Canada, despite previous use in Canadian research.\textsuperscript{24,25} We assessed the agreement of 32 nutrient intakes estimated from the FFQ with corresponding estimates from a second FFQ and the average of two 24-hour dietary recalls.

\section*{Methods}

\subsection*{Sample recruitment and study design}

During the autumn of 2003, random digit dialling methods were used to recruit a random population-based sample of Ontario women aged 25 to 74 years, stratified to represent the age distribution of women in Ontario. After excluding non-working or non-residential numbers and ineligible women (those who did not know English or were outside the age range), 235 women were identified, of whom 166 (71\%) agreed to participate. Data collection began in October 2003 and continued to February 2004.

Subjects were mailed the first FFQ (FFQ1), food portion guides, a short background questionnaire regarding age, marital status, ethnicity, education, height, weight and smoking history, and an availability form to assist study staff in contacting respondents for the interviews. Approximately 14 days after FFQ1 was completed, the first of two unannounced telephone-administered recalls was done, with a second one an average of 18 days later. About 26 days after the second recall, a second FFQ (FFQ2) was completed. The first and second FFQs were completed an average of 56 days apart. The study was approved by the Health Sciences Research Ethics Board at the University of Toronto.

\subsection*{Dietary assessment}

\textit{FFQ}

The 1998 version of the full-diet Block FFQ (Block98) is an eight-page, scannable, quantitative instrument that includes 109 food and beverage items (including alcohol); three multiple and nine single vitamin and mineral supplement items; additional questions to assess fats, fat-modified foods, types of milk and cereal; and summary cereal, fruit and vegetable questions. The FFQ food list was checked against other full-diet FFQs and summary cereal, fruit and vegetable questions. The assess fats, fat-modified foods, types of milk and cereal; mineral supplement items; additional questions to address alcohol); three multiple and nine single vitamin and mineral supplements. This approach to FFQ development was designed to yield 126 food and beverage items in the final analysis. The FFQ asked how often each food or beverage was usually consumed, and offered nine continuous responses ranging from ‘never’ to ‘every day’ for most foods. Four portion size choices were given using standard units (e.g. tablespoons or slices), or by referring to a separate sheet of photographs representing three-dimensional bowl and plate portions ranging from 1/4 to 2 cups. FFQs were self-administered by participants, and have been described as taking 30–40 min to complete (Torin Block, personal communication). Questionnaires were checked for completeness upon return, and telephone calls were made to collect missing information. FFQs were scanned and analysed by Block Dietary Data Systems using a nutrient content database based on the US Department of Agriculture (USDA) Nutrient Database for Standard Reference, national food consumption data (NHANES III and the Continuing Survey of Food Intakes by Individuals) and values from the published literature.\textsuperscript{1,5} Summary questions on the FFQ for cereals, fruit and vegetables were used to adjust nutrient intakes associated with these foods.\textsuperscript{15}

\subsection*{Dietary recalls}

Unannounced recalls, asking respondents to describe all foods, drinks and portion sizes consumed during the previous 24 h, were done over the telephone by one dietitian trained to use a Windows-based interview and nutrient analysis system, the Nutrient Data System for Research (NDS-R)\textsuperscript{26}, version 4.06 (2003, Nutrition Coordinating Center, University of Minnesota, MN, USA, Food and Nutrient Database 34). Training included attending a 2-day NDS-R workshop, and conducting 23 practice interviews evaluated by the study nutritionist for consistency and completeness.

A 16-page portion size booklet showing two-dimensional, actual-size food and drink models\textsuperscript{29} was mailed to respondents and used to guide the interviews.\textsuperscript{80} Standardised data collection was facilitated through NDS-R's scripted, multiple-pass approach,\textsuperscript{51} to which questions about alcohol and vitamin and mineral supplements were added to improve intake measurement,\textsuperscript{7,32,35} and comparability with the FFQ, where these items were also queried. When a respondent identified a supplement or food that was not included in NDS-R, a generic or best-fit alternative was used after matching ingredients identified by the respondent or, when unknown, by searching the Health Canada Drug Product Database\textsuperscript{54} or manufacturer websites. This best-fit approach was taken to improve the accuracy of intake reports because it was not known how thoroughly NDS-R included brands available in Canada. All interviews were checked for completeness and use of standardised methods by the study dietitian,
and a random sample of 10% were duplicate-checked by
the study nutritionist.

The two unannounced recalls included a weekday and
a weekend day for each respondent. Weekend days
included Friday to Sunday to capture food and alcohol
consumption patterns different from those on weekdays
(Monday to Thursday). It was not possible to
address seasonality within the time frame of this study.

Statistical analysis
All analyses were done using SAS, version 8.02 (SAS
statistics were computed to describe response rates,
demographic characteristics and average daily nutrient
intakes. Statistical analysis was restricted to 32 nutrients
common to the FFQ and recall databases, a list derived in
consultation with Block Dietary Data Systems and the Nutrition
Coordinating Center for NDS-R. Nutrient intakes
were reported in the same units by both databases, except
for alcohol, which was reported as grams of alcohol by
NDS-R but grams of alcoholic beverages by Block Dietary
Data Systems. Since we were unable to separate food from
supplements in the recalls, all micronutrient intake
include food and supplements combined. Energy adjustment
(density method) was done for major macronu-
trients, but not for micronutrients since food values
(associated with energy) could not be separated from
supplements values (not associated with energy).

Crude data from the FFQs and recalls were log-
transformed to improve normality, and outliers (intakes
greater than three standard deviations (> 3SD) from the
mean in FFQs or recalls) were removed before calculating
intake means, medians and Pearson correlation coeffi-
cients. On average, less than one outlier was excluded per
nutrient, ranging from zero for vitamin C (all methods) to
three for thiamine (recalls). One-way analysis of variance
was used to test for significant differences between mean
intakes. Reliability was assessed using Pearson correlation
coefficients between FFQ1 and FFQ2, while validity was
assessed using Pearson correlation coefficients between
FFQ1 and the average of two diet recalls. All
coefficients between FFQ1 and FFQ2, while validity was
used to test for significant differences between mean
intake means, medians and Pearson correlation coeffi-
cients. On average, less than one outlier was excluded per
nutrient, ranging from zero for vitamin C (all methods) to
three for thiamine (recalls). One-way analysis of variance
was used to test for significant differences between mean
intakes. Reliability was assessed using Pearson correlation
coefficients between FFQ1 and FFQ2, while validity was
assessed using Pearson correlation coefficients between
FFQ1 and the average of the two 24-hour recalls. All
validity coefficients were corrected for attenuation due to
random error in within-person variability, to allow a
reasonable estimate of true correlation, given that two
days of recall were collected.

Results
Of the 166 women who agreed to participate in this study,
115 (69%) completed FFQ1; of these, 105 (91%) completed the
first dietary recall, 96 (83%) completed the second
dietary recall and 93 (81%) completed FFQ2. Among those
women who completed both dietary recalls and one or
two FFQs (n = 96) (Table 1), the majority were married
(74%), Caucasian (82%), non-smokers (80%), with college
or university education (61%).

Mean, SD and median daily nutrient intakes for the two
FFQs and the average of two diet recalls are shown in
Table 2. Comparing FFQ1 and FFQ2, mean intakes were
similar across all nutrients except for energy and trans fatty
acids, where FFQ1 gave significantly higher estimates.
Mean intakes from FFQ1 and the recalls were similar for
total energy (1709 and 1844 kcal) although the FFQ gave
significantly lower estimates for carbohydrate, protein,
saturated fat and cholesterol, and higher estimates for total
fat (percentage of energy), trans fatty acids and
polyunsaturated fat. Mean intakes of 13 of the 17
micronutrients (food and supplements combined) were
not different between FFQ1 and the recalls, although the
FFQ gave significantly higher estimates for vitamin A, β-
carotene and vitamin C, and a lower estimate for selenium.

Pearson correlation coefficients between the two FFQs
(reliability) and between FFQ1 and the recalls (validity)
are presented in Table 3. Most correlations fell within the
range of 0.40 to 0.70, and most for macronutrients
increased when described in terms of energy density.
Correlation coefficients to estimate reliability between
FFQ1 and FFQ2 were found to be relatively high, with a
median of 0.75, and ranged from 0.57 (percentage energy
from protein) to 0.90 (percentage energy from alcohol) for
macronutrients, and from 0.65 (iron) to 0.88 (vitamin C)
for micronutrients.

All correlations to estimate validity increased with
deattenuation by 0.03 to 0.29. Deattenuated correlation

Table 1 Characteristics of women who completed two dietary recalls and at least FFQ1 in the Block98 FFQ validation study, 2003–2004 (n = 96)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Frequency, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>25–39</td>
<td>25 (26.0)</td>
</tr>
<tr>
<td>40–49</td>
<td>23 (24.0)</td>
</tr>
<tr>
<td>50–59</td>
<td>22 (22.9)</td>
</tr>
<tr>
<td>60+</td>
<td>26 (27.1)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
</tr>
<tr>
<td>Never married</td>
<td>11 (11.5)</td>
</tr>
<tr>
<td>Married/common law</td>
<td>71 (74.0)</td>
</tr>
<tr>
<td>Separated/divorced</td>
<td>10 (10.4)</td>
</tr>
<tr>
<td>Widowed</td>
<td>4 (4.2)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>79 (82.3)</td>
</tr>
<tr>
<td>Black</td>
<td>4 (4.2)</td>
</tr>
<tr>
<td>South Asian</td>
<td>2 (2.1)</td>
</tr>
<tr>
<td>Southeast Asian</td>
<td>2 (2.1)</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>1 (1.0)</td>
</tr>
<tr>
<td>Other</td>
<td>8 (8.3)</td>
</tr>
<tr>
<td>Highest education</td>
<td></td>
</tr>
<tr>
<td>Grade 1–8</td>
<td>5 (5.2)</td>
</tr>
<tr>
<td>Grade 9–13</td>
<td>23 (24.0)</td>
</tr>
<tr>
<td>Vocational school</td>
<td>9 (9.4)</td>
</tr>
<tr>
<td>College or university</td>
<td>49 (51.0)</td>
</tr>
<tr>
<td>Graduate degree</td>
<td>10 (10.4)</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>19 (19.8)</td>
</tr>
<tr>
<td>Former</td>
<td>30 (31.2)</td>
</tr>
<tr>
<td>Never</td>
<td>47 (49.0)</td>
</tr>
</tbody>
</table>

FFQ – food-frequency questionnaire.
coefficients between FFQ1 and recalls were generally moderate to high with an overall median of 0.59. For macronutrients, the median was 0.46, ranging from 0.11 (cholesterol) to 0.73 (percentage energy from carbohydrate); for micronutrients, the median was 0.65, ranging from 0.50 (iron) to 0.76 (folate). Only two of the 32 deattenuated correlation coefficients were less than 0.40, and both were associated with fats – cholesterol and monounsaturated fat.

Although reporting alcohol in different units hindered our ability to compare mean intakes from the FFQ and recalls, moderate to high reliability (0.86) and validity, especially when energy-adjusted (0.69), were observed for alcohol intake.

**Discussion**

Our estimates of reliability often exceeded the usual range for macronutrients (0.50 to 0.70)\(^2,3\) and were similar to those reported for micronutrients from food and supplements (0.53 to 0.83)\(^39,40\). Our estimates of validity were moderate for macronutrients compared with others (0.40 to 0.70)\(^3,22,23\), and moderate to high for micronutrients compared with those including food and supplements (0.21 to 0.83)\(^15,39–43\), and support the use of our FFQ in studies with Canadian women. Validity correlations less than 0.40 suggest FFQ shortcomings, based on our assumption that two recalls represent true intake.

**Comparison of average intakes with other studies**

Although ranking of individuals is the primary objective of most epidemiological studies\(^5,22\), nutrient intakes at the group level help assess how comparably each method describes the group mean\(^4^4\) and the comprehensiveness of the FFQ with respect to more detailed recalls\(^3,4^5\).

Our average macronutrient and energy intakes from recalls were comparable to those reported among...
women in recent validation studies in Canada\textsuperscript{26} and the USA\textsuperscript{4}, where energy intake hovered around 1700–1800 kcal. Average intakes from the FFQ were also similar for most macronutrients and energy compared with those estimated using the Block95 FFQ\textsuperscript{4}, but were sometimes lower than those using an FFQ developed for the Canadian Study of Lifestyle and Health\textsuperscript{26}, possibly due to different respondent and study factors. In contrast, our average micronutrient intakes were usually much higher in both the FFQ and recalls since neither of the other studies\textsuperscript{4,26} included supplements, although these may substantially increase intake over that from food alone\textsuperscript{7,46,47}. As an example, our average vitamin E intakes were at least twice those reported by others\textsuperscript{2,26}. We expect high supplement use in our sample: 25\% reported no supplements on both the FFQ and recalls since all subjects were female, and the majority were Caucasian, university-educated and non-smokers – all factors positively associated with supplement use\textsuperscript{46,48–51}. A review of our data showed the majority were supplement users: 25\% reported no supplements on both the FFQ and recalls. We were unable to separate food from supplements in the recall data to review their contributions separately; however, when we compared our results with those reporting combined food and supplement intakes among Caucasian women, our average intakes were usually similar\textsuperscript{45,46,47}.

Comparison of average intakes from FFQ1 and recalls: a focus on micronutrients

Mean intakes from FFQ1 and recalls were similar for most nutrients, although the FFQ gave different estimates for some macronutrients and higher estimates for micronutrients. Our lower FFQ estimates for protein, carbohydrate, cholesterol and saturated fats have been reported before, as have higher estimates of polyunsaturated fat and percentage energy from fat\textsuperscript{1,6,11,17–19,43}, and may result from limitations of the FFQ food list, response options or respondents to report diet accurately\textsuperscript{21,22,43}. Recent biomarker studies suggest that FFQs underestimate energy and protein more severely than recalls\textsuperscript{52–54}, and although this may not have resulted in a significant

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>FFQ1 and FFQ2 (n = 93)</th>
<th>FFQ1 and average of two 24-hour recalls (n = 96)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-deattenuated Pearson correlation coefficient (95% CI)</td>
<td>Non-deattenuated Pearson correlation coefficient (95% CI)</td>
</tr>
<tr>
<td>Energy (kcal)</td>
<td>0.74 (0.64, 0.82)</td>
<td>0.34 (0.15, 0.51)</td>
</tr>
<tr>
<td>Alcohol (g)†</td>
<td>0.86 (0.80, 0.91)</td>
<td>0.20 (0.10, 0.47)</td>
</tr>
<tr>
<td>Alcohol (% of energy)†</td>
<td>0.90 (0.85, 0.93)</td>
<td>0.40 (0.21, 0.55)</td>
</tr>
<tr>
<td>Carbohydrate (g)</td>
<td>0.80 (0.72, 0.87)</td>
<td>0.41 (0.22, 0.56)</td>
</tr>
<tr>
<td>Carbohydrate (% of energy)</td>
<td>0.78 (0.68, 0.85)</td>
<td>0.49 (0.32, 0.63)</td>
</tr>
<tr>
<td>Protein (g)</td>
<td>0.72 (0.61, 0.81)</td>
<td>0.30 (0.10, 0.47)</td>
</tr>
<tr>
<td>Protein (% of energy)</td>
<td>0.57 (0.41, 0.69)</td>
<td>0.29 (0.09, 0.46)</td>
</tr>
<tr>
<td>Fat (g)</td>
<td>0.72 (0.60, 0.81)</td>
<td>0.29 (0.10, 0.47)</td>
</tr>
<tr>
<td>Fat (% of energy)</td>
<td>0.70 (0.57, 0.79)</td>
<td>0.43 (0.25, 0.58)</td>
</tr>
<tr>
<td>Saturated fat (g)</td>
<td>0.72 (0.61, 0.81)</td>
<td>0.28 (0.09, 0.46)</td>
</tr>
<tr>
<td>Polyunsaturated fat (g)</td>
<td>0.72 (0.60, 0.81)</td>
<td>0.32 (0.13, 0.49)</td>
</tr>
<tr>
<td>Monounsaturated fat (g)</td>
<td>0.69 (0.57, 0.79)</td>
<td>0.24 (0.04, 0.42)</td>
</tr>
<tr>
<td>Trans fatty acids (g)</td>
<td>0.74 (0.63, 0.82)</td>
<td>0.37 (0.18, 0.53)</td>
</tr>
<tr>
<td>Cholesterol (mg)</td>
<td>0.69 (0.56, 0.78)</td>
<td>0.07 (–0.14, 0.27)</td>
</tr>
<tr>
<td>Dietary fibre (g)</td>
<td>0.84 (0.76, 0.89)</td>
<td>0.52 (0.36, 0.65)</td>
</tr>
<tr>
<td>Vitamin A (IU)‡</td>
<td>0.82 (0.74, 0.88)</td>
<td>0.40 (0.21, 0.55)</td>
</tr>
<tr>
<td>β-Carotene (µg)†</td>
<td>0.76 (0.65, 0.83)</td>
<td>0.27 (0.07, 0.44)</td>
</tr>
<tr>
<td>Thiamine (B1) (mg)‡</td>
<td>0.75 (0.65, 0.83)</td>
<td>0.68 (0.55, 0.77)</td>
</tr>
<tr>
<td>Riboflavin (B2) (mg)</td>
<td>0.75 (0.64, 0.83)</td>
<td>0.66 (0.53, 0.76)</td>
</tr>
<tr>
<td>Niacin (mg)‡</td>
<td>0.75 (0.64, 0.83)</td>
<td>0.64 (0.50, 0.74)</td>
</tr>
<tr>
<td>Vitamin B6 (mg)‡</td>
<td>0.79 (0.70, 0.86)</td>
<td>0.63 (0.49, 0.74)</td>
</tr>
<tr>
<td>Folate (µg)‡</td>
<td>0.78 (0.68, 0.85)</td>
<td>0.63 (0.50, 0.74)</td>
</tr>
<tr>
<td>Vitamin B12 (µg)‡</td>
<td>0.72 (0.60, 0.81)</td>
<td>0.55 (0.39, 0.68)</td>
</tr>
<tr>
<td>Vitamin C (mg)‡</td>
<td>0.88 (0.82, 0.92)</td>
<td>0.61 (0.46, 0.72)</td>
</tr>
<tr>
<td>Vitamin D (µg)†</td>
<td>0.76 (0.66, 0.83)</td>
<td>0.48 (0.31, 0.63)</td>
</tr>
<tr>
<td>Vitamin E (mg)†‡</td>
<td>0.81 (0.72, 0.87)</td>
<td>0.68 (0.56, 0.78)</td>
</tr>
<tr>
<td>Calcium (mg)</td>
<td>0.80 (0.71, 0.86)</td>
<td>0.56 (0.40, 0.68)</td>
</tr>
<tr>
<td>Copper (mg)†</td>
<td>0.83 (0.76, 0.89)</td>
<td>0.53 (0.36, 0.66)</td>
</tr>
<tr>
<td>Iron (mg)‡</td>
<td>0.69 (0.51, 0.75)</td>
<td>0.42 (0.23, 0.57)</td>
</tr>
<tr>
<td>Magnesium (mg)‡</td>
<td>0.82 (0.74, 0.88)</td>
<td>0.53 (0.37, 0.66)</td>
</tr>
<tr>
<td>Selenium (µg)‡</td>
<td>0.69 (0.57, 0.79)</td>
<td>0.41 (0.23, 0.56)</td>
</tr>
<tr>
<td>Zinc (mg)‡</td>
<td>0.75 (0.64, 0.83)</td>
<td>0.51 (0.34, 0.64)</td>
</tr>
</tbody>
</table>

CI – confidence interval; FFQ – food-frequency questionnaire.
*All data are based on crude intakes with log-transformed outliers removed.
†Based on grams of alcohol in recalls, and grams of alcoholic beverages in FFQ.
‡Total intake based on supplements and food.
difference in total energy intake, it may help explain our lower FFQ macronutrient estimates.

FFQs that include a large number of items to represent specific food groups encourage higher estimates of intake\(^{6,8,11,21,55,59,60}\). Since vitamin A, \(\beta\)-carotene and vitamin C are associated with fruit and vegetable consumption\(^{57,59}\), higher estimates for these micronutrients may be partially due to the large number of fruit and vegetable items on our FFQ: the Block98 FFQ included 28 fruit or vegetable foods among its original 109 items.

Higher FFQ estimates of vitamins C and A and \(\beta\)-carotene have been reported previously with the Block FFQ compared with recalls or records, both among women\(^{1,4,18,20}\) and men\(^{1,14}\) and whether using earlier versions of the full-diet\(^{1,4,14,18}\) or reduced FFQ\(^{59,60}\). This is not, however, unique to the Block FFQ, and has been seen with other FFQs\(^{21,26,39,43,59}\).

Fruit and vegetable adjustment questions were introduced to help correct this effect\(^{15}\), but we and others\(^{20}\) observed higher estimates despite adjustment, suggesting either incomplete correction or the role of other factors, such as the inclusion of supplements or limitations of the reference method. Higher Block FFQ estimates had been observed whether intakes were based on food only\(^{1,4,14,18}\), food with adjustment\(^{18}\), food and supplements\(^{18}\), or food and supplements with adjustment (our study).

Although the inclusion of supplements improves intake measurement and classification\(^{7,42}\), their contribution may be so great relative to food\(^{7,32,33,46}\) that any correction based on food alone may be masked, particularly when other sources of dietary assessment error are not adequately reduced. It is notable that another study, which included food and supplements with adjustment, did not show higher FFQ estimates compared with records\(^{15}\). This may be partly explained by the inclusion of more days in the reference method and hence a better assessment of usual diet and supplement use than in our study, affecting both individual and group estimates\(^{18}\). Block et al. collected three 4-day records per participant\(^{15}\) compared with our two 24-hour recalls. Although one day may adequately capture average group energy intake in a sample of 100 women, many more days would be needed for the more variable nutrients from food, such as vitamin A\(^{60}\). More than four days, and perhaps as many as 14, may be needed to adequately identify supplement use\(^{42}\).

**Correlation of intakes from FFQ1 and recalls (validity)**

How well the FFQ captures usual diet affects average intakes and correlations between the FFQ and recalls. Under our validity assumptions, FFQ limitations mentioned earlier account for low correlation coefficients. However, since no dietary assessment method is error-free\(^{21,22,61}\), our reference method based on two recalls must also be examined.

For nutrients with small within-person to between-person ratios, only a few days of recall are needed to capture usual intake, whereas those with large ratios require many more days, especially when evaluating individual rather than group intake\(^{22,57,62–65}\). Among macronutrients, fat variables have the highest ratios, especially cholesterol; of the micronutrients, vitamin A and \(\beta\)-carotene have the highest ratios\(^{57,60,64,65}\). It is not surprising that correlation coefficients lower than 0.40 were associated with two of the most variable macronutrients – cholesterol (0.11) and monounsaturated fat (0.35). Additionally, although two days of recall may be adequate to estimate validity, correlations between FFQ and recalls have been shown to increase with the number of recall days\(^{15,22,40}\).

Comparing our deattenuated correlation coefficients with those found in other studies among women\(^{4,26}\), correlations for energy were comparable. Those for many macronutrients, however, were higher than ours (most notably for cholesterol: 0.11 vs. 0.57 and 0.55), although less so when using an earlier version of our FFQ\(^{4}\). This trend was reversed, however, for micronutrients, where we observed higher correlation coefficients, sometimes dramatically so for those commonly consumed as supplements\(^{5,46,47,49,51}\), such as vitamin E (0.71 vs. 0.37 and 0.28). Since the inclusion of supplements improves micronutrient correlations, we also compared our results with the few studies among Caucasian women that included food and supplements combined. Many of our correlations continued to be higher, particularly for B vitamins, as did our average micronutrient correlation (0.65 vs. 0.58)\(^{15,40,41}\).

**Discrepancy between validity correlations for macronutrients and micronutrients**

The discrepancy between lower correlations associated with macronutrients but higher correlations associated with micronutrients may be explained by shortcomings and strengths of our FFQ, as well as our study. Many of the confidence intervals around correlations were wide, especially for macronutrients, due to our small sample size, and this somewhat limited the interpretation of our findings. Additionally, usual diet and the full extent of dietary variation may not have been captured by two days of recall. Correlations may have improved with additional days of recall\(^{15,22,40}\), as in Jain et al. (three days)\(^{26}\) and Subar et al. (four days)\(^{4}\), and over different seasons\(^{4}\). This shortcoming in our recall data would affect macronutrients, especially the most variable ones such as fats. A supplementary analysis highlights this limitation and its effects (data not shown). When grams of food consumed daily were examined, 18% of our recall data fell within the top 2.5% of the consumption
distribution in a large nationally representative sample of Americans (NHANES 1999–2000), suggesting unusually high intake reports over our two days of recall. Some of these high reports may be due to misunderstandings of the two-dimensional portion size measures used for recalls. For one respondent, estimates using a picture of a glass instead of a tablespoon added up to 40 tablespoons of cream in coffee; for another, the portion size picture chosen for meat in one dish added to 0.57 kg of ground beef. These high reports would have been minimally affected by removing less than 1% of our data as outliers. When these high reports (18%) were removed from the analysis, deattenuated correlations for most fat variables increased from less than or close to 0.40 to 0.50 or greater.

In contrast, validity correlations associated with all micronutrients, including the most variable, were higher in our study than in those based on more days of recall but which excluded supplement data. Our correlations were also higher than those reported in studies including supplement data. Four possible explanations arise within the context of our sample of women and their supplement use. First, given the important contribution of supplements relative to food, the addition of supplement data not only improves micronutrient intake measurements at the group and individual levels but also correlations between different methods, even when food intake data may be limited. Second, although supplement use is common among women in our study, there may be considerable intra-individual heterogeneity in use that improved correlations. Third, by adding questions to the recall interviews about supplement intake, we encouraged reports similar to those prompted by the FFQ list of supplements. (Similarly, the addition of alcohol recall questions may have contributed to the moderate to high alcohol correlations.) Fourth, the Block98 FFQ included modifications which may have encouraged higher micronutrient correlations than those seen with the Block95 FFQ. Modifications included the incorporation of additional vitamin supplements, seasonal and canned fruit, an additional frequency option (a few times a year), different portion options, and a portion guide whose three-dimensional representations may have offered cognitive advantages to estimating intake.

Other factors may have affected validity although the direction of their effects is uncertain. Although both NDS-R and the Block Dietary Data Systems databases are based on the USDA Nutrient Database for Standard Reference, differences in inclusions and accuracy may affect recall and FFQ analyses, and comparisons between them. Since neither database was modified for Canadian foods, the current validation study was facilitated by basing analyses on US data only. However, future Canadian studies using the Block98 FFQ and its affiliated database may need to be interpreted cautiously, particularly when examining foods or nutrients affected by fortification regulations that differ between Canada and the USA, such as for folate, vitamin D, calcium, zinc, iron or vitamin A. Similarly, it is uncertain how the different FFQ and recall food portion models affected the accuracy and comparability of intake reports, and how this might have affected validity.

While our study sample was similar to the general population in terms of including visible minorities, it differed substantially on education. Although our recruitment of university-educated respondents (61%) likely assisted FFQ and recall completion and validity, this selection bias may hinder the generalisability of our findings to the broader population of Canadians, where only 33% are so highly educated. Participant characteristics may have also affected our energy estimates, since factors such as body mass index (BMI) are associated with energy underreporting. In a group of Canadian women where 35% were low energy reporters (LERs) by FFQ, mean energy intake among non-LER women was higher than ours. Since our sample had a higher mean BMI (26.4 vs. 23.6 kg m⁻², data not shown) this may have increased the extent of underreporting in our group. Although energy intakes from our FFQ1 and recalls were similar, both are likely underestimates given that underreporting is extensive in all methods, particularly FFQs. Although macronutrients are reported differently by LERs and non-LERs, it is not known to what extent each macronutrient contributes to energy underestimation, how well different dietary methods capture this, or how this affects validity.

An additional issue, and one whose effects cannot be assessed here, is the suggestion that traditional validation studies are flawed by systematic intake-related and person-specific biases (e.g. overreporting of healthy foods) which, by being large and shared by FFQs and report-based reference measures such as recalls, undermine critical validation assumptions of unbiasedness and independence. Using various FFQs from the European Prospective Investigation into Cancer and Nutrition (EPIC), including a modified Willett FFQ used in the UK arm, models correcting for these biases have suggested that validation studies traditionally overestimate FFQ accuracy and validity, and modify the ability to detect diet–disease relationships, although it is not clear how this would specifically alter our results. Future validation studies may benefit from the inclusion of biomarker reference methods such as urinary nitrogen, potassium or doubly labelled water to better evaluate measurement error associated with protein, potassium or energy, respectively, and its impact on the design and interpretation of epidemiological studies where FFQs are applied.
Validation of the Block98 FFQ

Acknowledgements

This research was supported by the Canadian Breast Cancer Research Alliance with special funding support of the Canadian Breast Cancer Foundation no. Ontario Chapter (CBCRA grant no. 013572).

The authors wish to thank Sarah Winterton for assistance with dietary assessment and study co-ordination.

References


Palaniappan U, Starkey LJ, O’Loughlin JO, Gray-Donald K. Fruit and vegetable consumption is lower and saturated fat intake is higher among Canadians reporting smoking. *Journal of Nutrition* 2001; 131: 1952–8.


Basiotis PP, Welsh SO, Cronin FJ, Kelsay JL, Mertz W. Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. *Journal of Nutrition* 1987; 117: 1608–11.


74 Livingstone MBE, Black AE. Markers of the validity of reported energy intake. *Journal of Nutrition* 2003; 133: 895S–920S.

