Published in final edited form as:

J Acad Nutr Diet. 2014 April ; 114(4): 613–621. doi:10.1016/j.jand.2013.11.017.

Evaluation of Web-Based, Self-Administered, Graphical Food Frequency Questionnaire

Alan R. Kristal, Dr. P.H.^{1,2}, Ann S. Kolar, MPH, RD¹, James L. Fisher, PhD³, Jesse J. Plascak³, Phyllis J. Stumbo, PhD⁴, Rick Weiss, MS⁵, and Electra D. Paskett, Ph.D.^{3,6}

Alan R. Kristal: akristal@fhcrc.org; Ann S. Kolar: kolar.ann@gmail.com; James L. Fisher: Jay.Fisher@osumc.edu; Phyllis J. Stumbo: phyllis-stumbo@uiowa.edu; Rick Weiss: Weiss@viocare.com; Electra D. Paskett: Electra.Paskett@osumc.edu

^{1,2}Member and Associate Head, Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Department of Epidemiology, University of Washington, P.O. Box 19024, 1100 Fairview Ave N (M4-B402), Seattle, WA 98109, Phone: 206-667-4686, Fax: 206-667-7850

¹Cancer Prevention Program, Fred Hutchinson Cancer Research Center, P.O. Box 19024, 1100 Fairview Avenue N (M4-B402), Seattle, WA 98109, Phone: 425-603-0090, Fax: 425-603-1121

³Research Scientist, Comprehensive Cancer Center and James Cancer Hospital, The Ohio State University, 1590 N. High St. Suite 525, Columbus, Ohio 43201, Phone: 614.293.964, Fax: 614.293.6667

⁴Institute for Clinical and Translational Sciences, University of Iowa, Research Nutritionist Emeritus, 1700 S. First Avenue - Suite 17, Iowa City, IA 52240, Phone: 319-530-7840, Fax: 319-248-0222

⁵Princeton University MS Electrical Engineering and Computer Science, President, Founder, Chief Wellness Engineer, Viocare, Inc., 145 Witherspoon Street, Princeton, NJ 08542, Phone: 609-497-4600 x10, Fax: 609-497-0660

^{3,6}*Marion N. Rowley Professor of Cancer Research and Director, Division of Cancer Prevention and Control, Department of Internal Medicine, College of Medicine, Professor, Division of Epidemiology, College of Public Health, Associate Director for Population Sciences, Comprehensive Cancer Center, Ohio State University, Suite 525, 1590 North High Street, Columbus, OH 43201, Phone: (614) 293-3917, Fax: (614) 293-5611*

Abstract

Computer-administered food frequency questionnaires (FFQs) can address limitations inherent in paper questionnaires, by allowing very complex skip patterns, portion size estimation based on food pictures and real-time error checking. This manuscript evaluates a web-based FFQ, the Graphical Food Frequency System (GraFFS). Participants completed the GraFFS, six, telephone-administered 24-hr dietary recalls over the next 12 weeks, followed by a second GraFFS. Participants were 40 men and 34 women, ages 18–69, living in the Columbus, OH area. Intakes of energy, macronutrients and 17 micronutrients/food components were estimated from the GraFFS

© 2014 Academy of Nutrition and Dietetics. Published by Elsevier Inc. All rights reserved.

Corresponding Author: Alan R. Kristal, Dr.P.H., Cancer Prevention Program, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N (M4-B402), Seattle, WA 98109, akristal@fhcrc.org.

Institutional Review Board approval:

This research was approved by the Ohio State University Institutional Review Board

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

and the mean of all recalls. Bias (recalls minus the second GraFFS) was -9%, -5%, +4% and -4% for energy and percentages of energy from fat, carbohydrate and protein. De-attenuated, energy-adjusted correlations (inter-method reliability) between the recalls and the second GraFFS for fat, carbohydrate, protein and alcohol were 0.82, 0.79, 0.67 and 0.90; for micronutrients/food components the median was 0.61 and ranged from 0.40 for zinc to 0.92 for β -carotene. The correlations between the two administrations of the GraFFS (test-retest reliability) for fat, carbohydrate, protein and alcohol were 0.60, 0.63, 0.73 and 0.87; among micronutrients/food components the median was 0.67 and ranged from 0.49 for vitamin B12 to 0.82 for fiber. The measurement characteristics of the GraFFS were at least as good as those reported for most paper FFQs, and its high inter-method reliability suggests that further development of computer-administered FFQs is warranted.

Keywords

Food Frequency Questionnaire; Dietary Assessment; Computer-Administered Questionnaires; Evaluation; Reliability

INTRODUCTION

The Food Frequency Questionnaire (FFQ) has been widely used in epidemiological and clinical research as a means to assess dietary intake. A substantial amount of dietary assessment research has focused on the structure, content and analysis of FFQs, with the purpose of improving their overall validity and reliability as measures of usual food intake. For example, based on cognitive research¹, Thompson et al² showed that first asking about a food group (e.g., milk) followed by questions on types used (low- vs. regular-fat) was superior to asking about each type of food separately. Another example is work by Patterson et al,³ which demonstrated the importance of modifying FFQs for populations with unique dietary patterns, such as older US women who consume a wide variety of fat-modified foods. Nevertheless, recent studies based on objective measures of nutrient intake have found that FFQs are very poor measures of both total energy⁴ and some micronutrients, such as selenium and vitamin E, which are dispersed across many foods with concentrations that vary widely based on their origin and formulation.⁵⁻¹⁰ While it is important to understand the limitations of FFQs, they are useful for measuring intakes of nutrients concentrated in relatively few foods (e.g., calcium from dairy products),¹¹ they can be used to measure aspects of food use, such as consumption of fruits and vegetables,¹² that are related to chronic disease risk, and they can be used to formulate personalized feedback in clinical interventions to promote healthful dietary change.^{13,14} Thus, further research to improve FFQs for both clinical and research applications is well motivated.

One inherent limitation to most FFQs is that they are paper-based forms. Thus, errors such as skipped questions or multiple marks are common, and incorporating complex skip patterns, a broad and varying number of portions size options, and extensive food and portion-size graphics is challenging. Furthermore, data from paper forms must be entered into analysis software, which makes it unfeasible to provide real-time feedback in a clinical setting. Computer-administered questionnaires offer straightforward solutions to these limitations, and there are several examples of computer-administered FFQs in the published literature.¹⁵ In their most simple application, computer-administered FFQs match paper questionnaires; this allows the flexibility of using either a paper or computerized questionnaire interchangeably, but the benefits from computer administration are limited to direct data entry, real-time error checking and rapid analysis.^{16,17} More complex computer-administered FFQs allow the use of complex skip algorithms to make questionnaires less burdensome and the flow of questions more logical,¹⁸ and several have added food

photographs to assist with portion size estimation.^{19–22} Of these questionnaires, only those from NutritionQuest,¹⁷ the National Cancer Institute¹⁸ and Viocare²² are publically available, and only the National Cancer Institute questionnaire has been formally evaluated.²⁰

Here we give the results of a study to evaluate the measurement characteristics of the VioScreen²² Graphical Food Frequency System (GraFFS). The GraFFS is an algorithm-driven, computer-delivered dietary questionnaire, which was developed to address three primary limitations of paper FFQs: (1) respondent errors, such as skipped questions, missing pages and multiple marks, are common in paper questionnaires; these are eliminated in the GraFFS because complete answers are required before moving to the next question; (2) very complex skip patterns, which can minimize participant burden and support a logical flow of questions, are unwieldy to implement in paper forms; these are easily implemented in the GraFFS; and (3) portion sizes on paper forms are generally restricted to “small,” “medium” and “large,” and defined as 50%, 100% and 150% of a standard portion described for each food, which does not reflect the wide range of portions commonly consumed; in the GraFFS up to six portions sizes are given both as pictures and in text. The underlying hypothesis guiding and funding the development of the GraFFS is that it would provide dietary intake data that were superior to paper FFQs for clinical and research use.

METHODS

The GraFFS is based on paper FFQs developed at the Fred Hutchinson Cancer Research Center.²³ These questionnaires are similar in length and design, with minor variations in portion size options and food definitions, and have been used in large, epidemiological studies and intervention trials, including the Women’s Health Initiative (WHI),²⁴ the Prostate Cancer Prevention Trial,^{25,26} the Selenium and Vitamin E Cancer Prevention Trial,²⁷ and the Vitamin and Lifestyle Cohort Study.²⁸ The GraFFS collects data on the use of 156 food items or food groups, along with portion size based on between 3 and 6 pictures for each food. The GraFFS also collects information on preparation method and food formulations (e.g., calcium-fortified, fat-reduced), as well as use of fats and condiments added during preparation or at the table. The structure of the GraFFS is designed to minimize participant burden in two ways. First, at the beginning of each section (e.g., beverages, fruits, vegetables), the user is presented a screen of illustrated thumbnails corresponding to all foods in that section; users select foods and beverages that they consume at least once a month and only the selected foods/food groups are subsequently queried. Second, questions used to refine nutrient estimates, such as added sugar, formulation or preparation method, are only asked when required. Examples of GraFFS screens can be found at www.viocare.com/vioscreen.aspx.

The goal of this study was to measure the validity and reliability of nutrient intake generated by the GraFFS. The issues surrounding validation of dietary measurement tools are both complex and controversial. The most important consideration is the criterion or “true” measure that is used to assess validity. Optimal criterion measures are recovery biomarkers, however there are only two: doubly-labeled water can be used to measure absolute energy intake and urinary nitrogen excretion can be used to measure protein intake.⁴ Concentration biomarkers, such as blood micronutrient concentrations, can also be used as criterion measures. However, because there are many determinants of serum micronutrient concentrations, they are more useful as an unbiased standard for comparing two alternative dietary assessment approaches (e.g., FFQs vs. food diaries) than as criterion measures themselves.⁹ Lastly, a more rigorous approach to measuring food intake, such as weighed food records or repeated 24-hour recalls, can serve as a criterion measure; however, because all self-report measures are subject to participant-specific biases,²⁹ measurement errors in

two instruments are correlated and validity estimates will be inflated.³⁰ This evaluation study used six, interviewer-administered 24-hr recalls as the criterion measure, which is a commonly used approach in the scientific literature.³¹ Repeated dietary recalls provide a better measure of most nutrients than a FFQ, but they do not generate “true” values of nutrient intake. A better description of the agreement between the dietary recalls and the GraFFS is therefore termed “inter-method reliability,” which is used throughout this report.³² In addition to inter-method reliability, we also report bias, which is the average difference between the two instruments, and test-retest reliability, which is the correlation between repeated administrations of the GraFFS.

Participants were recruited in late 2009 through study posters placed around The Ohio State University (OSU) campus and in the Columbus, Ohio area and by advertisements placed in the OSU Daily News. Recruitment targeted men and women who were between 18 and 70 years of age and lived in the Columbus area; initial eligibility criteria assessed during a telephone interview included willingness to complete all study procedures and age. Eligible and willing participants traveled to the OSU Comprehensive Cancer Center (OSU-CCC) to complete informed consent and the first administration of the GraFFS. The research was approved by Institution Review Board at the Ohio State University.

The study procedures included an initial administration of the GraFFS, six, 24-hour recalls administered by telephone over the subsequent three months, and a final administration of the GraFFS. Both administrations of the GraFFS were completed at the OSU-CCC, and the referent time period for reporting usual food use was the past 3 months. A short questionnaire at the beginning of the study collected demographic and general information on household responsibility for food selection, shopping and preparation; a questionnaire at the end of the study collected information on participants’ evaluation of the GraFFS. Participants were given a printed report on their nutrient intake based on the GraFFS and a \$50 gift certificate to a local sports store, as partial compensation for their time and travel expenses.

Telephone-administered 24-hr recalls were conducted by the Nutrition Assessment Shared Resource at the Fred Hutchinson Cancer Research Center.²³ Interviewers complete a 3-week training on the use of the Nutrition Data System for Research (NDS-R)³³ plus additional training on protocols for administering 24-hr recalls. Calls are monitored for quality control, and all recalls are reviewed by a registered dietitian. The dietary recall interviews were not pre-scheduled with participants, preventing knowledge of when their diet would be recorded, and were administered on various days of the week including two weekend days. The interview consisted of three parts: 1) listing what the participant ate and drank in the previous day; 2) describing each food, its preparation method, and all beverages in detail; and 3) reviewing the foods consumed to probe for missing items and additional intake information. Participants used a portion size booklet to help describe portions to the interviewer (<http://sharedresources.fhcrc.org/content/sample-serving-size-booklet>). All data were entered directly into a computer using the Nutrition Data System for Research (NDS-R) software, developed by the Nutrition Coordinating Center at the University of Minnesota, Minneapolis, MN.³³ Nutrient analyses were completed using the NDS-R 2009 database, which was also used to develop the nutrient database for analysis of the GraFFS.³⁴

A total of 109 persons were screened for eligibility by telephone, of whom 105 were eligible and 100 attended the first study visit. Using a set of *a-priori* criteria commonly applied to FFQs, 18 participants who could not adequately complete the first GraFFS (2 men and 6 women reported consuming less than 1000 kilocalories; 7 men and 1 woman reported consuming more than 4500 kilocalories; and 1 man and 1 woman reported consuming fewer than 25 different foods) were excluded, leaving 82 eligible participants to proceed to

completing the 24-hr recall interviews. Additional exclusions included 5 participants who completed fewer than four, 24-hour recalls and 3 participants who did not complete the final GraFFS administration. Data presented here are for the 74 participants who completed a baseline and follow-up GraFFS and at least four 24-hour recalls.

Analyses were completed for nutrients and food components that are of general interest in health research and clinical dietetics. All nutrient values, with the exception of macronutrients expressed as percentage of total energy, were log transformed before analysis. The final administration of the GraFFS was used for the analyses of inter-method reliability and bias, because its time reference (in the past 3 months) covered the period in which the 24-hr recalls were administered. The mean of the 24-hr recalls was used as the comparison measure of dietary intake.

Bias was calculated as the mean difference between the GraFFS and the 24-hr recalls. These differences were normally distributed and thus analyses are based on untransformed data. The inter-method reliability was calculated as the Pearson and Spearman correlation coefficients between the GraFFS and the 24-hr recalls. These are given unadjusted, adjusted for energy using the residual method,³⁵ and further adjusted for random error in the 24-hr recalls. The energy-adjusted measure is commonly used in the evaluation of food frequency questionnaires, because it corrects for a tendency of respondents to either over- or under-report their frequencies of consumption. Adjustment for random error in the 24-hr recalls used the following formula:

$$p_c = p_o \sqrt{1 + [\sigma_w^2 / (\sigma_b^2 / n)]}$$

Where p_o is the observed correlation between the energy-adjusted nutrients from the GraFFS and recalls and σ_w^2 is the within-person variation, σ_b^2 is the between-person variation and n is the number of replicate measurements.³⁶ This adjusts the inter-method reliability to the value that would be observed if 24-hr recalls were collected each day. This computation is necessary when comparing across inter-method reliability studies that use different numbers of recall days, because the magnitude of the inter-method reliability is a function of the reliability of the criterion measure, which increases as more recall days are used to calculate its mean. The energy- and error-adjusted values from the evaluation of the WHI FFQ²⁴ are given for comparison, because the design of the WHI FFQ is similar to the GraFFS and its evaluation design and statistical methods match the approach used here. With the exception of the biostatisticians (JLF and JJP), the authors of this report had no access to the study data used in this report.

RESULTS and DISCUSSION

Demographic characteristics of the study population are given in Table 1. The sample included similar numbers of men and women, of whom 63% were under age 40 years; 36.5% of participants were non-white, the overwhelming majority of whom were African-American. More than 80% had completed college, over half were married or living as married, and more than 80% reported having some or most of the household responsibility for meal planning, food shopping and cooking.

Table 2 gives data related to bias, including the geometric mean values of nutrient intakes estimated from the 24-hr recalls and both administrations of the GraFFS, as well the specific measure of bias (24hr recall – end-of-study GraFFS). Mean intakes of macronutrients and most micronutrients tended to be lower in the end of study compared to baseline GraFFS administration. The GraFFS underestimated total energy, fat, carbohydrate and protein by

9%, 15%, 5% and 12%, respectively (Bias percentages not shown in Table). However, when macronutrients were expressed as percentage of total energy, the GraFFS underestimated fat and protein by 5% and 4%, respectively, and overestimated carbohydrate by 4%. The GraFFS tended to overestimate intakes of micronutrients. For micronutrients such as β -carotene, genestein and daidzein, which are concentrated in specific foods and eaten sporadically, bias was over +50%. Bias was approximately +35% for vitamins C, D and B12, and under 10% for niacin, thiamin, riboflavin, vitamin B6, folate and iron.

Table 3 gives results for the inter-method reliability (agreement) of the GraFFS compared to 24-hr recalls. The unadjusted Pearson correlations ranged from a high of 0.75 for alcohol to 0.39 for energy; for micronutrients they ranged between 0.66 for β -carotene to 0.42 for genestein. Energy adjustment increased the correlations for macronutrients, substantially so for fats and carbohydrate. Energy adjustment had mixed effects on micronutrients, with the largest increase for calcium and largest decrease for B-12. As expected, further adjustment of error in the 24-hr recalls increased correlations for all measures. Energy-adjusted, de-attenuated inter-method reliabilities for macronutrients ranged from 0.61 for percentage energy from polyunsaturated fat to 0.90 for alcohol; for micronutrients they ranged from 0.40 for zinc to 0.92 for β -carotene. There were striking differences between these results for the GraFFS and those published from the evaluation of the WHI FFQ.²⁴ Of the macronutrients, only the values for alcohol were similar; for all others, correlations from the GraFFS were substantially higher. The comparison across micronutrients was more mixed; GraFFS values for iron, zinc, vitamin D, and thiamin were modestly lower, while those for β -carotene, retinol, and vitamins B12 and C were substantially higher than from the WHI Initiative FFQ. Inter-method reliability was also similar to or higher than those reported for other paper FFQs used in major epidemiological studies.³⁷

Table 4 gives results for test-retest reliability, as the correlations between the two GraFFS administrations. Note that the administrations were separated by at least 3 months and the time reference for each was “over the past 3 months,” so they did not assess the same time period and differences may be influenced by seasonal food availability and dietary change. Nevertheless, correlations were good for macronutrients, with a low of 0.59 for polyunsaturated fat and a high of 0.87 for alcohol. Correlations for micronutrients tended to be somewhat lower, ranging from 0.49 for Vitamin B12 to 0.74 for riboflavin.

Overall, based on the 12-item exit questionnaire, participants rated the GraFFS highly. More than 80% reported that the pictures helped them to select their portions sizes and only 20% judged the questionnaire length to be too long. Of note were the 98% who would fill out the questionnaire if asked by their doctors and the 100% who felt the system was easy to use. The least favorable aspect was the GraFFS was on “It was easy to select how often I ate each food,” which was rated neutral by 16.2% and disagree by 1.2%.

Comparisons between the GraFFS and the two other publically-available computer-administered instruments are not straightforward. The NutritionQuest FFQ matches a popular paper-based questionnaire developed by Block et al.¹⁷ The most recent published evaluation of a NutritionQuest paper FFQ (Block98) was completed by Boucher et al.,³⁸ who compared this FFQ to the mean of two 24hr dietary recalls. The energy-adjusted, de-attenuated inter-method reliabilities for macronutrients were 0.41, 0.51, 0.41, 0.41, 0.35 and 0.42 for protein, carbohydrate, total, saturated, monounsaturated and polyunsaturated fats, which are all substantially lower than those reported here for the GraFFS. The intra-method reliability of micronutrients from the Block98 questionnaire cannot be compared to those for the GraFFS, because the Block98 evaluation included micronutrient intake from dietary supplements. There are no published comparisons of the paper and computer-administered Block98 questionnaires, and therefore we cannot know whether computer-administration would

substantially improve these results. The US National Cancer Institute computer-administered Diet History Questionnaire (DHQ) incorporates cognitive research in the structure and order of questionnaire items and uses complex algorithms to reduce response burden: participants first select foods eaten at least once a month, and subsequent questions focus on these foods only. Beasley et al.²⁰ administered both the paper and computer-administered (with portion size pictures) versions of the DHQ, and compared these to nutrient intakes estimated from both a 4-day food record and the mean of two 24-hr recalls. For all macro and most micronutrients, correlations of the computer-administered version with records and recalls were higher than those with the paper version. The inter-method reliability of energy-adjusted but not de-attenuated macronutrients from the computer-administered DHQ were considerably lower than those reported here for GraFFS: using the food record as the criterion measure, correlations were 0.40, 0.30, 0.39, 0.56, 0.36 and 0.21 for protein, carbohydrate, total, saturated, monounsaturated and polyunsaturated fats, respectively; using the two 24-hr recalls as the criterion measure, these correlations were 0.45, 0.38, 0.30, 0.29, 0.21 and 0.10. The inter-method reliabilities for most micronutrients measured by the GraFFS were slightly higher than those for the DHQ, however differences were modest. In part, the lower intra-method reliability of macronutrients measured by the NCI questionnaire could be attributable to the evaluation study design: the evaluation of the DHQ used different nutrient databases for the FFQ, food records and 24-hr recalls; only 4 days (food records) or 2 days (24-hr recalls) of food intake were used as the criterion measure of usual nutrient intake; and the time period referenced for the FFQ was “over the past year” but the food record and recalls were completed over 4 weeks. However, because the GraFFS collects much more data on food additions and formulations than the DHQ, and it uses up to 6 portion size options rather than the “small,” “medium” and “large” used on DHQ, the GraFFS may also be capturing data on food intake more accurately.

There are several limitations of this study. First, we did not compare the GraFFS to either a paper FFQ or a different computer-administered questionnaire within the same study population; rather we based our comparisons on previously published studies. Differences in study characteristics such as eligibility criteria and participant age, sex and education make comparison across studies evaluating single instruments difficult. For example, we excluded participants who, based on their first administration of the GraFFS, could not satisfactorily complete the questionnaire, but this approach of excluding participants who reported physiologically-impossible, long-term energy intakes is common across most studies. Other studies used small samples of volunteers from existing cohorts²⁴ or persons willing to comply with very demanding study protocols,³⁹ and any other of these selection criteria could bias results to be different than would be otherwise obtained.⁴⁰ Our population was highly educated, which may limit the generalizability of this study to other populations. Second, we did not have unbiased nutritional biomarkers as criterion measures. Our primary comparison was with the WHI paper FFQ for several reasons. Their design and the foods included were similar, both used the same system for developing a nutrient database from the NDS-R system,²³ and both used the same study design and statistical approach for evaluation. Thus, we judge this to be a reasonable comparison of the GraFFS to a paper FFQ. We note, however, that we could not test whether differences between the GraFFS and WHI questionnaires were statistically significant, because the necessary data from the WHI FFQ were not available; given that the confidence intervals around de-attenuated correlations are larger than those for raw correlations, it is unlikely that any contrasts reached standard levels of statistical significance. Our secondary comparison was with the NCI computer-administered DHQ: this questionnaire is somewhat comparable to the GraFFS, but because the design and analysis of its evaluation differed substantially any inferences about the relative validity of the computer-administered DHQ and GraFFS are at best speculative. Further research to evaluate the GraFFS should incorporate the type of design used by Subar and colleagues⁴¹ in which 3 commonly-used FFQs were compared

within a single study sample; new studies should also include biomarkers of nutrient intake as additional criterion measures that are not subject to recall or response-set bias.^{30,42,43} Lastly, the mean reported intake of total energy and all macronutrients was lower on the second compared to first administration of the GraFFS. Whether this reflects true differences due, for example, to seasonable availability of foods, or a training effect from completing multiple 24-hr recalls is uncertain. Further research is needed to better understand how repeat administration of the GraFFS affects its measurement characteristics.

From a practical perspective, we describe several applications of the GraFFS, which is commercialized under the name VioScreen, to illustrate how computerized FFQs can be used as a research and clinical tool. Researchers have incorporated it into pharmaceutical trials and observational studies in which data on dietary patterns were needed to address study aims. In many of these applications, the GraFFS was modified to include additional foods and food preparation methods to better capture food patterns of interest. Several large hospitals and individual clinical nutritionists have incorporated the GraFFS into their protocols for patient care. Examples of these applications are that GraFFS provided nutritionists information on foods contributing most to intakes of saturated fat, oxalate and rapidly absorbed carbohydrate, which was then used to formulate specific targets for dietary behavior change. The GraFFS has been incorporated into corporate wellness programs, in which computer-generated feedback on dietary patterns is returned to participants. Lastly, the GraFFS is being used in medical and undergraduate schools as part of their nutrition curriculum.

Conclusion

The measurement characteristics of the GraFFS are, overall, at least as good as those generally reported for paper FFQs. One unexpected finding was the high inter-method reliability of macronutrients measured by the GraFFS, which suggests that addressing limitations of paper-administered questionnaires could improve the accuracy of FFQs. We cannot know from this study whether and to what extent the use multiple portion size pictures, the incorporation of complex skip patterns, or the elimination of missing or non-interpretable responses contributed to the higher inter-method reliability of macronutrients measured by the GraFFS compared to other computer-administered and many paper FFQs. Further research should address this question, because it would provide the basis for further improvement of computer-administered dietary questionnaires.

Acknowledgments

This study was funded by the National Institutes of Health, National Cancer Institute, Grant number CA097560

References

1. Smith AF. Cognitive processing in long-term dietary recall. *Vital and Health Statistics*. 1991; 6:1–34. [PubMed: 1688237]
2. Thompson FE, Subar AF, Brown CC, et al. Cognitive research enhances accuracy of food frequency questionnaire reports: results of an experimental validation study. *J Am Diet Assoc*. Feb; 2002 102(2):212–225. [PubMed: 11846115]
3. Patterson RE, Kristal AR, Coates RJ, et al. Low-fat diet practices of older women: prevalence and implications for dietary assessment. *J Am Diet Assoc*. 1996; 96(7):670–679. [PubMed: 8675910]
4. Schatzkin A, Kipnis V, Carroll RJ, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) Study. *Int J Epidemiol*. 2003; 32(6):1054–1062. [PubMed: 14681273]
5. Levander OA. The need for measures of selenium status. *J Am Coll Toxicol*. 1986; 5(1):37–44.

6. Stryker WS, Kaplan LA, Stein EA, Stampfer MJ, Sober A, Willett WC. The relation of diet, cigarette smoking, and alcohol consumption to plasma beta-carotene and alpha-tocopherol levels. *Am J Epidemiol.* 1988; 127(2):283–296. [PubMed: 3257350]
7. Hunter DJ, Morris JS, Chute CG, et al. Predictors of selenium concentration in human toenails. *Am J Epidemiol.* 1990; 132(1):114–122. [PubMed: 2356804]
8. Brunner E, Stallone D, Janeja M, Bingham S, Marmot M. Dietary assessment in Whitehall II: comparison of 7 d diet diary and food-frequency questionnaire and validity against biomarkers. *Br J Nutr.* 2001; 86(3):405–414. [PubMed: 11570993]
9. Dixon LB, Subar AF, Wideroff L, Thompson FE, Kahle LL, Potischman N. Carotenoid and tocopherol estimates from the NCI Diet History Questionnaire are valid compared with multiple recalls and serum biomarkers. *J Nutr.* 2006; 136(12):3054–3061. [PubMed: 17116719]
10. Talegawkar SA, Johnson EJ, Carithers T, Taylor HA, Bogle ML, Tucker KL. Total α -tocopherol intakes are associated with serum α -tocopherol concentrations in African American adults. *J Nutr.* 2007; 137(10):2297–2303. [PubMed: 17885014]
11. Musgrave KO, Giambalvo L, Leclerc HL, Cook RA, Rosen CJ. Validation of a quantitative food frequency questionnaire for rapid assessment of dietary calcium intake. *J Am Diet Assoc.* Oct; 1989 89(10):1484–1488. [PubMed: 2794308]
12. Kristal AR, Vizenor NC, Patterson RE, Neuhouser ML, Shattuck AL, McLerran D. Precision and bias of food frequency-based measures of fruit and vegetable intakes. *Cancer Epidemiol Biomarkers Prev.* 2000; 9:939–944. [PubMed: 11008912]
13. Kristal AR, Curry SJ, Shattuck AL, Feng Z, Li S. A randomized trial of a tailored, self-help dietary intervention: The Puget Sound Eating Patterns Study. *Prev Med.* 2000; 31:380–389. [PubMed: 11006063]
14. Maes L, Cook TL, Ottovaere C, et al. Pilot evaluation of the HELENA (Healthy Lifestyle in Europe by Nutrition in Adolescence) Food-O-Meter, a computer-tailored nutrition advice for adolescents: a study in six European cities. *Public Health Nutr.* 2011; 14(07):1292–1302. [PubMed: 21281542]
15. Falomir Z, Arregui M, Madueno F, Corella D, Coltell O. Automation of Food Questionnaires in Medical Studies: a state-of-the-art review and future prospects. *Computers in biology and medicine.* Oct; 2012 42(10):964–974. [PubMed: 22898338]
16. Matthys C, Pynaert I, De Keyzer W, De Henauw S. Validity and reproducibility of an adolescent web-based food frequency questionnaire. *J Am Diet Assoc.* 2007; 107(4):605–610. [PubMed: 17383266]
17. NutritionQuest. [accessed Feb 2012] Block Questionnaires. <http://www.nutritionquest.com> [Error! Hyperlink reference not valid.. <http://www.nutritionquest.com>
18. NCI. National Cancer Institute. Risk Factor Monitoring and Methods, Diet History QuesTionnaire (DHQ). National Cancer Institute; <http://riskfactor.cancer.gov/DHQ> [accessed Feb 2012.]
19. Labonte ME, Cyr A, Baril-Gravel L, Royer MM, Lamarche B. Validity and reproducibility of a web-based, self-administered food frequency questionnaire. *Eur J Clin Nutr.* Feb; 2012 66(2):166–173. [PubMed: 21934698]
20. Beasley JM, Davis A, Riley WT. Evaluation of a web-based, pictorial diet history questionnaire. *Public Health Nutr.* May; 2009 12(5):651–659. [PubMed: 18547450]
21. Vereecken CA, De Bourdeaudhuij I, Maes L. The HELENA online food frequency questionnaire: reproducibility and comparison with four 24-h recalls in Belgian-Flemish adolescents. *Eur J Clin Nutr.* May; 2010 64(5):541–548. [PubMed: 20216566]
22. Viocare. [accessed Feb 2012. accessed Feb 2012.] VioScreen. <http://www.viocare.com><http://www.viocare.com>
23. Nutrition. Nutrition Assessment Shared Resource. the Fred Hutchinson Cancer Research Center; http://www.fhcrc.org/science/shared_resources/nutrition
24. Patterson RE, Kristal AR, Tinker LF, Carter RA, Bolton MP, Agurs-Collins T. Measurement characteristics of the Women's Health Initiative food frequency questionnaire. *Ann Epidemiol.* 1999; 9(3):178–187. [PubMed: 10192650]
25. Thompson IM, Coltman CA, Crowley J. Chemoprevention of prostate cancer: the Prostate Cancer Prevention Trial. *Prostate.* 1997; 33(3):217–221. [PubMed: 9365551]

26. Kristal AR, Arnold KB, Schenk JM, et al. Dietary patterns, supplement use and risk of symptomatic benign prostatic hyperplasia (BPH): Results from the Prostate Cancer Prevention Trial. *Am J Epidemiol.* 2008; 167(8):925–934. [PubMed: 18263602]
27. Lippman SM, Klein EA, Goodman PJ, et al. Effect of selenium and vitamin E on risk of prostate cancer and other cancers. The Selenium and Vitamin E Cancer Prevention Trial (SELECT). *JAMA.* 2009; 301(1):39–51. [PubMed: 19066370]
28. White E, Patterson RE, Kristal AR, et al. VITamins And Lifestyle Cohort Study: study design and characteristics of supplement users. *Am J Epidemiol.* 2004; 159:83–93. [PubMed: 14693663]
29. Neuhouwer ML, Tinker L, Shaw PA, et al. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *Am J Epidemiol.* 2008; 167(10): 1247–1259. [PubMed: 18344516]
30. Kipnis V, Midthune D, Freedman LS, et al. Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am J Epidemiol.* 2001; 153:394–403. [PubMed: 11207158]
31. Willett, W. *Nutritional Epidemiology*. 2. New York: Oxford University Press; 1998. Reproducibility and Validity of Food-Frequency Questionnaires; p. 101-147.
32. White, E.; Armstrong, BK.; Saracci, R. *Principles of exposure measurement in epidemiology*. 2. New York: Oxford University Press; 2008.
33. Schakel SF. Maintaining a nutrient database in a changing marketplace: Keeping pace with changing food products-A research perspective = Maintenance d'une base de donnée nutritionnelle dans un marché changeant: Evoluer avec les produits alimentaires changeant- perspective de recherche. *J Food Comp Anal.* 2001; 14(3):315–322.
34. Kristal, AR.; Shattuck, AL.; Williams, AE. Food frequency questionnaires for diet intervention research. Paper presented at: 17th National Nutrient Databank Conference; 1992; Baltimore, MD.
35. Willett, WC. *Nutritional Epidemiology*. 2. New York: Oxford University Press; 1998. Correction of effects of measurement error; p. 302-320.
36. Beaton GH, Milner J, Corey P, et al. Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. *Am J Clin Nutr.* Dec; 1979 32(12): 2546–2559. [PubMed: 506977]
37. Thompson FE, Kipnis V, Midthune D, et al. Performance of a food-frequency questionnaire in the US NIH-AARP (National Institutes of Health-American Association of Retired Persons) Diet and Health Study. *Public Health Nutr.* Feb; 2008 11(2):183–195. [PubMed: 17610761]
38. Boucher B, Cotterchio M, Kreiger N, Nadalin V, Block T, Block G. Validity and reliability of the Block98 food-frequency questionnaire in a sample of Canadian women. *Public Health Nutr.* Feb; 2006 9(1):84–93. [PubMed: 16480538]
39. Thompson FE, Midthune D, Subar AF, Kipnis V, Kahle LL, Schatzkin A. Development and evaluation of a short instrument to estimate usual dietary intake of percentage energy from fat. *J Am Diet Assoc.* 2007; 107(5):760–767. [PubMed: 17467371]
40. Marks GC, Hughes MC, van der Pols JC. Relative validity of food intake estimates using a food frequency questionnaire is associated with sex, age, and other personal characteristics. *J Nutr.* Feb; 2006 136(2):459–465. [PubMed: 16424128]
41. Subar A, Thompson F, Kipnis V, et al. Comparative validation of the Block, Willett, and the National Cancer Institute food frequency questionnaires: The Eating and America's Table study. *Am J Epidemiol.* 2001; 152:1089–1099. [PubMed: 11744511]
42. Kristal AR, Andrilla CH, Koepsell TD, Diehr PH, Cheadle A. Dietary assessment instruments are susceptible to intervention-associated response set bias. *J Am Diet Assoc.* Jan; 1998 98(1):40–43. [PubMed: 9434649]
43. Miller TM, Abdel-Maksoud MF, Crane LA, Marcus AC, Byers TE. Effects of social approval bias on self-reported fruit and vegetable consumption. *Nutr J.* 2008; 7(118)

Table 1

Demographic Characteristics of Study Sample

Characteristic	(n=74)	
	N	%
Sex		
Male	40	54.1%
Female	34	46.0%
Age		
18 – 29	30	40.5%
30 – 39	17	23.0%
40 – 49	8	10.8%
50 – 59	9	12.2%
60 – 69	10	13.5%
Race/Ethnicity		
White	47	63.5%
African American	20	27.0%
Hispanic	2	2.7%
Other	5	6.8%
Refused	0	0.0%
Marital Status		
Single	24	32.4%
Married, Living as Married	41	55.4%
Divorced/Separated/Widowed	9	12.2%
Refused	0	0.0%
Education, years		
<= 12	2	2.7%
13 – 15	11	14.9%
16	34	45.9%
17+	27	36.5%
Refused	0	0.0%

Household Responsibilities	Most	Some	Little	Don't eat at home
Meal Planning	55.4%	33.8%	10.8%	0.0%
Food shopping	58.1%	31.1%	10.8%	0.0%
Cooking	50.0%	33.8%	16.2%	0.0%

Table 2

Mean intakes of selected nutrients from six, 24-hour Recalls and the GraFFS, and Bias in the GraFFS Compared to 24-hr Recalls (n = 74)

Nutrient	24-hr Recalls		GraFFS		Bias ^d Mean ^c (SE), p-value ^d
	Mean ^b (95 Percent CI)		Baseline Mean ^b (95 Percent CI)	End of Study Mean ^b (95 Percent CI)	
Energy kcal	1967 (1274, 3011)		1959 (1086, 3533)	1720 (846, 3498)	-182 (73.4), <0.02
Fat – Total					
g	70.1 (38.9, 126.5)		67.4 (31.5, 144.0)	58.6 (26.8, 127.7)	-10.3 (3.0), <0.001
%en ^c	32.7 (21.3, 44.1)		31.7 (20.5, 42.9)	31.2 (19.4, 43.0)	-1.6 (0.6), <0.008
Saturated					
g	22.4 (11.0, 45.6)		21.3 (9.0, 50.4)	18.5 (7.8, 43.8)	-3.5 (1.0), <0.001
%en ^c	10.6 (5.1, 16.1)		10.1 (5.4, 14.8)	9.9 (5.2, 14.6)	-0.6 (0.3), <0.05
Polyunsaturated					
g	16.3 (8.3, 31.8)		14.4 (6.5, 32.1)	12.7 (5.4, 30.0)	-3.3 (0.8), <0.009
%en ^c	7.7 (4.0, 11.4)		6.9 (3.6, 10.2)	6.8 (3.5, 10.1)	-0.9 (0.2), <0.001
Monounsaturated					
g	25.3 (13.5, 47.5)		25.8 (11.4, 58.6)	22.2 (10.2, 48.4)	-2.8 (1.1), <0.02
%en ^c	11.8 (7.1, 16.5)		12.2 (6.5, 17.9)	11.9 (5.8, 18.0)	0.0 (0.3), <0.03
EPA and DHA					
g	0.13 (0, 0.48)		0.17 (0, 0.54)	0.19 (0, 0.82)	0.1 (0.0), <0.93
Carbohydrate					
g	242.3 (142.6, 411.6)		247.2 (131.6, 464.1)	221.4 (101.5, 483.0)	-12.6 (9.7), <0.20
%en ^c	50.0 (36.1, 63.9)		51.2 (35.9, 66.5)	52.0 (37.3, 66.7)	2.0 (0.7), <0.004
Protein					
g	83.9 (48.4, 145.5)		80.6 (38.9, 167.3)	70.8 (30.6, 164.0)	-10.0 (4.2), <0.02
%en ^c	17.4 (9.8, 25.0)		16.9 (10.2, 23.6)	16.7 (10.0, 23.4)	-0.7 (0.4), <0.09
Alcohol					
g	2.1 (0, 31.1)		3.5 (0, 47.4)	3.1 (0, 39.9)	1.4 (0.7), <0.06
%en ^c	2.8 (0, 8.1)		3.6 (0, 9.9)	3.8 (0, 11.1)	1.0 (0.3), <0.002

Nutrient	24-hr Recalls Mean ^b (95 Percent CI)	GraFFS		Bias ^d Mean ^c (SE), p-value ^d
		Baseline Mean ^b (95 Percent CI)	End of Study Mean ^b (95 Percent CI)	
Dietary fiber, g	20.5 (9.8, 42.9)	21.3 (10.0, 45.6)	19.1 (7.9, 46.1)	-0.1 (0.8), <0.25
Retinol, mcg RE	1054 (428, 2592)	1353 (459, 3984)	1224 (358, 4188)	309 (97), <0.002
Beta-carotene, mcg, RE	2724 (544, 13,630)	4064 (728, 22,697)	3429 (599, 19,930)	1515 (486), <0.002
Vitamin C, mg	87.4 (27.9, 273.1)	113.7 (36.2, 354.2)	113.3 (35.5, 361.4)	30.4 (9.2), <0.001
Vitamin D, mcg	408 (125, 1326)	523 (114, 1901)	459 (83, 2540)	141 (42), <0.001
Niacin, mg NE	41.3 (24.3, 70.1)	39.3 (18.4, 83.9)	35.2 (15.2, 81.5)	-4.3 (2.1), <0.05
Thiamin, mg	17.6 (10.2, 30.6)	18.0 (8.1, 40.0)	16.4 (6.7, 40.4)	-0.3 (1.1), <0.81
Riboflavin, mg	21.8 (11.4, 41.7)	24.0 (9.8, 59.1)	22.0 (8.2, 58.6)	1.5 (1.2), <0.22
Vitamin B6, mg	21.3 (11.1, 40.9)	22.2 (9.4, 52.5)	21.1 (8.2, 54.1)	1.1 (1.4), <0.43
Vitamin B12, mcg	48.4 (17.1, 137.0)	62.8 (22.9, 175.9)	59.1 (15.6, 223.6)	17.3 (5.5), <0.002
Folate, mcg FE	579 (296, 1130)	590 (226, 1541)	539 (194, 1495)	-0.5 (37.5), <0.87
Calcium, mg	907 (454, 1808)	1153 (441, 3011)	1012 (365, 2807)	175 (47.8), <0.001
Iron, mg	16.1 (8.9, 29.1)	17.5 (7.4, 41.3)	15.0 (5.6, 40.0)	0.1 (1.0), <0.94
Zinc, mg	11.1 (6.3, 19.7)	14.0 (6.2, 31.8)	12.6 (5.1, 30.9)	2.3 (0.8), <0.005
Genistein, mg	590 (24.3, 14,328)	665 (27.7, 15,994)	837 (27.7, 25,336)	1597 (667), <0.02
Daidzein, mg	488 (19.3, 12,333)	545 (20.7, 14,328)	699 (24.5, 19,930)	1150 (521), <0.03

^aEnd of study GraFFS – 24-hour recall

^bGeometric mean, back-transformed for ease of interpretation, except where noted

^cArithmetic mean

^dP-value for test that bias = 0.

Table 3

Inter-method reliability of selected nutrients measured by the GraFFS compared to six, 24-hour dietary recalls (n = 74)

Nutrient ^a	Unadjusted ^b		Energy-adjusted ^b		De-Attenuated Pearson	
	Pearson	Spearman	Pearson	Spearman	GraFFS	WHI ^c
Energy, kcal	0.39	0.44				
Fat, total fat						
g	0.42	0.40	0.63	0.53	0.82	0.64
%en ^d	0.59	0.54	-	-	0.75	-
Saturated						
g	0.44	0.44	0.68	0.64	0.84	0.63
	0.66	0.64	-	-	0.82	-
Polyunsaturated						
g	0.40	0.35	0.44	0.39	0.67	0.54
%en ^d	0.40	0.39	-	-	0.61	-
Monounsaturated						
g	0.47	0.45	0.63	0.55	0.85	0.64
%en ^d	0.55	0.55	-	-	0.74	-
EPA and DHA						
g	0.68	0.55	0.62	0.46	0.89	0.59
Carbohydrate						
g	0.47	0.47	0.67	0.63	0.79	0.67
%en ^d	0.67	0.66	-	-	0.82	-
Protein						
g	0.41	0.44	0.54	0.51	0.67	0.51
%en ^d	0.53	0.49	-	-	0.62	-
Alcohol						
g	0.75	0.74	0.73	0.68	0.90	0.89
%en ^d	0.68	0.74	-	-	0.90	-

Nutrient ^a	Unadjusted ^b		Energy-adjusted ^b		De-Attenuated Pearson	
	Pearson	Spearman	Pearson	Spearman	GraFFS	WHI ^c
Dietary fiber, g	0.58	0.61	0.75	0.73	0.85	0.70
Retinol, mcg RE	0.51	0.48	0.59	0.55	0.81	0.30
Beta-carotene, mcg RE	0.66	0.64	0.69	0.69	0.92	0.52
Vitamin C, mg	0.46	0.42	0.52	0.51	0.66	0.37
Vitamin D, mcg	0.51	0.58	0.43	0.40	0.56	0.70
Niacin, mg NE	0.47	0.52	0.48	0.42	0.61	0.63
Thiamin, mg	0.47	0.52	0.39	0.38	0.53	0.66
Riboflavin, mg	0.56	0.62	0.60	0.61	0.71	0.65
Vitamin B6, mg	0.44	0.39	0.49	0.46	0.69	0.66
Vitamin B12, mcg	0.53	0.48	0.38	0.27 ^e	0.53	0.21
Folate, mcg FE	0.49	0.51	0.43	0.41	0.61	0.57
Calcium, mg	0.52	0.52	0.61	0.64	0.79	0.73
Iron, mg	0.47	0.47	0.36	0.38	0.50	0.66
Zinc, mg	0.43	0.46	0.30 ^e	0.25 ^f	0.40	0.47
Genistein, mg	0.42	0.38	0.47	0.42	0.57	-
Daidzein, mg	0.51	0.50	0.55	0.52	0.65	-

^aLog-transformed, except as noted.

^bAll correlations significantly different from zero (p<0.001), except where noted.

^cRef 24, Table 4.

^dNot transformed

^ep<0.01

^fp<0.05

Table 4

Reliability of selected nutrients measured by the GraFFS at Baseline and 12-Week (n=74)

Nutrient	Pearson Correlation ^a
Energy, kcal ^b	-
Fat, total	
g	0.60
%en ^c	0.60
Saturated	
g	0.68
%en ^c	0.73
Polyunsaturated	
g	0.59
%en ^c	0.58
Monounsaturated	
g	0.61
%en ^c	0.59
EPA and DHA	
g	0.52
Carbohydrate	
g	0.63
%en ^c	0.65
Protein	
g	0.73
%en ^c	0.72
Alcohol	
g	0.87
%en ^c	0.72
Dietary fiber, g	0.82
Retinol, mcg RE	0.60
Beta-carotene, mcg RE	0.72
Vitamin C, mg	0.64
Vitamin D, mcg	0.52
Niacin, mg NE	0.70
Thiamin, mg	0.70
Riboflavin, mg	0.74
Vitamin B6, mg	0.72
Vitamin B12, mcg	0.49
Folate, mcg FE	0.67

Nutrient	Pearson Correlation ^a
Calcium, mg	0.61
Iron, mg	0.72
Zinc, mg	0.69
Genistein, mg	0.55
Daidzein, mg	0.63

^aLog-transformed, energy-adjusted values, except as noted.

^bLog-transformed only

^cNot transformed or energy-adjusted