



PRACTICE OF EPIDEMIOLOGY

Is Shorter Always Better? Relative Importance of Questionnaire Length and Cognitive Ease on Response Rates and Data Quality for Two Dietary Questionnaires

Amy F. Subar,¹ Regina G. Ziegler,² Frances E. Thompson,¹ Christine Cole Johnson,³ Joel L. Weissfeld,⁴ Douglas Reding,⁵ Katherine H. Kavounis,⁶ and Richard B. Hayes⁷ for the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial Investigators

In this study, the authors sought to determine the effects of length and clarity on response rates and data quality for two food frequency questionnaires (FFQs): the newly developed 36-page Diet History Questionnaire (DHQ), designed to be cognitively easier for respondents, and a 16-page FFQ developed earlier for the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. The PLCO Trial is a 23-year randomized controlled clinical trial begun in 1992. The sample for this substudy, which was conducted from January to April of 1998, consisted of 900 control and 450 screened PLCO participants aged 55–74 years. Controls received either the DHQ or the PLCO FFQ by mail. Screenees, who had previously completed the PLCO FFQ at baseline, were administered the DHQ. Among controls, the response rate for both FFQs was 82%. Average amounts of time needed by controls to complete the DHQ and the PLCO FFQ were 68 minutes and 39 minutes, respectively. Percentages of missing or uninterpretable responses were similar between instruments for questions on frequency of intake but were approximately 3 and 9 percentage points lower ($p \leq 0.001$) in the DHQ for questions on portion size and use of vitamin/mineral supplements, respectively. Among screenees, response rates for the DHQ and the PLCO FFQ were 84% and 89%, respectively, and analyses of questions on portion size and supplement use showed few differences. These data indicated that the shorter FFQ was not better from the perspective of response rate and data quality, and that clarity and ease of administration may compensate for questionnaire length. *Am J Epidemiol* 2001;153:404–9.

diet; diet surveys; epidemiologic methods; food habits; nutrition assessment; nutrition surveys; questionnaires

In epidemiologic and public health research, cooperation in completing questionnaires is elicited in a variety of ways. Depending on the research questions, study design, funding, and practical considerations, questionnaires can be

interviewer- or self-administered, long or short, easy or complex. Efforts to maximize response rates generally include some combination of introductions, reminders, and follow-ups through in-person, telephone, or mail contacts. Incentives such as payment, gifts, or medical information may also enhance final response rates. Response rates, in turn, are important measures of study success, validity, and generalizability.

Following telephone data collection methods, mail surveys are the most common method of obtaining questionnaire data (1, 2). They are economical and practical in many research situations and are common in large-scale epidemiologic research. A review of past research indicates that for mailed surveys, follow-up is the most important factor in increasing response rates, followed by prior notice, financial incentives, special mailing, a positive perception of the survey sponsors, provision of a postage-paid return envelope, and personalization (2). It is less clear how questionnaire length affects response rates.

Nutritional epidemiologic studies frequently incorporate mailed, self-administered questionnaires to collect dietary data, particularly in large prospective studies or intervention trials in which participants are spread over a wide geo-

Received for publication January 31, 2000, and accepted for publication May 26, 2000.

Abbreviations: DHQ, Diet History Questionnaire; FFQ, food frequency questionnaire; PLCO, Prostate, Lung, Colorectal, and Ovarian [Cancer Screening Trial].

¹Risk Factor Monitoring and Methods Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD.

²Epidemiology and Biostatistics Program, Division of Cancer Etiology and Genetics, National Cancer Institute, Bethesda, MD.

³Josephine Ford Cancer Center, Detroit, MI.

⁴Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA.

⁵Marshfield Medical Research and Education Foundation, Marshfield, WI.

⁶Westat, Incorporated, Rockville, MD.

⁷Environmental Epidemiology Branch, Division of Cancer Etiology and Genetics, National Cancer Institute, Bethesda, MD.

Reprint requests to Dr. Amy F. Subar, National Cancer Institute, Executive Plaza North, Room 4005, 6130 Executive Blvd., MSC 7344, Bethesda, MD 20892-7344 (e-mail: amy_subar@nih.gov).

graphic area. Many such studies rely on the use of machine-readable food frequency questionnaires (FFQs), which generally require a substantial time commitment (30–60 minutes) from participants. Several studies have addressed various methodological issues related to response rates and data quality specifically for mailed, self-administered FFQs (3–8). Similar to findings for mailed surveys in general, factors such as preliminary notification, personal contact, incentives, and follow-up contacts increase response rates for FFQs (3–5), but findings regarding the effect of questionnaire length and design on response rates or data quality have been inconsistent (3, 6–8).

Although data conflict regarding whether questionnaire length actually decreases response rates, concerns about this and increased respondent burden can preclude or limit detailed dietary assessment in nutritional epidemiologic studies. In this analysis, we examined response rates and data quality between two dietary questionnaires: a new 36-page FFQ designed to be cognitively easier for respondents and a 16-page FFQ. The longer questionnaire incorporated many elements of the “total design method” for mail surveys developed by Dillman (1, 2), the purpose of which is to maximize response rates and data quality. The method attends in detail to visible and cognitive aspects of the questionnaire development and survey implementation process, including the ordering of questions, graphic design, and content.

MATERIALS AND METHODS

Study design

The Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, begun in 1992, is a 23-year randomized controlled clinical trial of screening procedures for prostate, lung, colon, and ovarian cancer involving 148,000 men and women aged 55–74 years. The 10 screening centers for the trial are the Georgetown University Medical Center (Washington, DC), the Henry Ford Health System (Detroit, Michigan), the Marshfield Medical Education and Research Foundation (Marshfield, Wisconsin), the Pacific Health Research Institute (Honolulu, Hawaii), the University of Alabama (Birmingham, Alabama), the University of Colorado (Denver, Colorado), the University of Minnesota (Minneapolis, Minnesota), the University of Pittsburgh (Pittsburgh, Pennsylvania), the University of Utah (Salt Lake City, Utah), and Washington University (St. Louis, Missouri). Participants have been randomized into either the screened arm or the control arm of the trial. Details regarding the study’s design and purpose are reported elsewhere (9).

Additional etiologic studies of cancer were developed in both the screening and control components of the trial. Diet was one of a number of exposures to be assessed. All participants randomized into the screening arm of the trial received a PLCO FFQ (described below) at baseline either by mail or in person (depending on the practice specific to each site).

The current substudy, conducted from January to April of 1998, resulted from an effort to determine the practical utility of adding a recently developed FFQ, the Diet

History Questionnaire (DHQ) (described below), to both the control and the screening arms of the PLCO Trial. The new instrument is more detailed and cognitively based than the original PLCO FFQ, and therefore we thought that it might be advantageous to include it in the study. Testing of the willingness of both screenees and controls to complete the instrument was considered essential in deciding whether or not to add the DHQ to the study protocol. Three screening centers were selected to participate in this analysis: the Henry Ford Health System, the University of Pittsburgh, and the Marshfield Medical Education and Research Foundation. At each center, 300 control participants (a total of 900) were randomly selected; half were randomized to receive, by mail, the DHQ and half to receive the PLCO FFQ. The mailing, conducted by each center, was accompanied by a cover letter and a postage-paid return envelope. For participants who did not return their questionnaires within 3 weeks, up to five telephone calls were made by staff at each center (at different times and days of the week) in an effort to make a single follow-up contact. A second questionnaire was provided to participants who reported losing or misplacing the first questionnaire. In addition to either the DHQ or the PLCO FFQ, participants were given a one-page questionnaire on which they recorded the amount of time they had needed to complete the questionnaire. The participants were asked to answer two qualitative questions: 1) “Which of the following categories best describes how much help you needed in answering how often you eat foods or your usual portion size?” (response categories: no help needed, little help needed, moderate help needed, much help needed) and 2) “How easy or difficult was the dietary questionnaire to fill out?” (response categories: very easy, somewhat easy, somewhat difficult, very difficult).

In the screening arm of the study, 150 participants at each of the three screening centers (a total of 450), who had previously received a PLCO FFQ at baseline as part of the trial protocol, were selected sequentially as their year 3 clinic visits occurred to receive the DHQ, to be completed either on-site or at home (with a postage-paid return envelope). Whichever procedures each screening center had previously used for administration and follow-up of the PLCO FFQ at baseline were adhered to for the DHQ administration in year 3. Screened participants were not asked the questions about length and difficulty of the dietary instrument, because there were no comparable data from the administration of the PLCO FFQ at baseline.

Food frequency questionnaires

In 1994, investigators at the National Cancer Institute began to develop a new, self-administered, machine-readable FFQ in the hope of improving the FFQs currently being used in epidemiologic research. The intent was to improve both the cognitive and database aspects of FFQs. To this end, cognitive testing was conducted (10) which resulted in the creation of the DHQ, a 36-page FFQ that was improved with respect to content, order, wording, and

layout. The DHQ asked about frequency of intake for 114 individual food items; for 109 of the items, the DHQ asked about portion size by providing a choice of three portion size ranges. For 47 of these food items, 1–7 additional embedded questions asked about related aspects of consumption, such as seasonal intake, food type (e.g., low-fat, lean, diet, caffeine-free), fat additions, or fat type. The DHQ also asked 1) five questions about the proportion of the time the respondent used various types of margarine, mayonnaise, sour cream, cream cheese, and salad dressing; 2) four summary questions; and 3) nine questions on the use of vitamin and mineral supplements, six of which queried about frequency of intake and duration and five of which queried about usual dose. A copy of the DHQ can be found at <http://www-dccps.ims.nci.nih.gov/ARP/DHQ>.

In 1993, prior to completion of the DHQ in 1998, a shorter FFQ that incorporated some of the DHQ innovations was developed for the PLCO Trial. This PLCO FFQ is a 16-page machine-readable booklet with the appearance of a standard Block/National Cancer Institute or Willett grid-format FFQ. It was developed to characterize total diet while minimizing participant burden. The PLCO FFQ contains questions on 137 individual food items, 60 of which include a query about usual portion size; 10 questions about meat cooking practices; and 14 questions about intake of vitamin/mineral supplements, 12 of which contain detailed queries regarding current use, dosage, duration, and use 2 and 5 years previously. Overall, the PLCO FFQ included many fewer items overall and many more items per page than the DHQ. A complete copy of the PLCO FFQ can be found at <http://dcp.nci.nih.gov/plco/diet>.

Analyses

Differences between the two questionnaires for each arm of the trial were assessed by comparing response rates, percentages of unusable responses, and, for the control arm, times needed to complete the questionnaire and answers to qualitative questions. Absolute numbers for questions on frequency of intake, portion size, and supplement use differed between the two FFQs; therefore, for those three types of questions, the mean percentages of questionnaire responses that were missing or uninterpretable (because of stray marks or double-responding found in the machine scanning) were compared using *t* tests among controls and paired *t* tests among screenees. We conducted χ^2 tests to determine differences in response rates and in answers to qualitative questions between the DHQ and PLCO FFQ. A *t* test was used to determine differences between DHQ and PLCO FFQ respondents in the amount of time taken to complete the instruments.

RESULTS

The study sample of controls and screenees combined was 51 percent female and 49 percent male. Fourteen percent of participants were aged 50–59 years, 54 percent were aged 60–69 years, and 32 percent were aged 70 years or more. Approximately 11 percent had less than a high school

education, 29 percent had graduated from high school, 33 percent had some post-high-school education, and 27 percent were college graduates or had postcollege education. Randomization of the controls to receive either the DHQ or the PLCO FFQ showed that the participants were evenly distributed by gender and were within 3 percentage points by age group.

Table 1 shows the percentages of DHQs and PLCO FFQs returned by controls and screenees. Among controls, the response rates for the two questionnaires for all centers combined were nearly identical (approximately 82 percent). Within centers, response rates for controls by FFQ type differed by 0.7–8.3 percentage points. None of the differences were statistically significant. In the screenee group, the percentage of participants who returned the DHQ was 5 percentage points lower than the PLCO FFQ return rate for these same participants 3 years earlier (84 percent vs. 89 percent) ($p \leq 0.05$). Furthermore, the DHQ return rate by center was consistently lower than the earlier return rate for the PLCO FFQ.

After initial response rates were determined, a decision was made to exclude respondents for whom 33 percent of frequency-of-intake responses were missing or uninterpretable (at or above the 99th percentile of the frequency distribution). Four of the DHQs and five of the PLCO FFQs returned by controls were excluded, as were five DHQs and no PLCO FFQs returned by screenees. Finally, because the analyses for screenees required within-subject comparisons, 21 persons were excluded because they had not successfully completed both the DHQ and the PLCO FFQ by the above exclusion criterion. This left 357 screenees for the final analyses.

The mean amounts of time taken by controls to complete the DHQ and the PLCO FFQ were 68 minutes and 39 minutes, respectively. The times varied little by center and were significantly different overall and for each center ($p \leq 0.0001$).

TABLE 1. Percentages of food frequency questionnaires returned by control and screened participants in a substudy of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, 1998

	% of questionnaires returned		
	No. of participants	Diet History Questionnaire	PLCO food frequency questionnaire
Controls			
All centers	900	81.9	81.6
Detroit, MI	300	72.3	75.0
Pittsburgh, PA	300	84.0	78.7
Marshfield, WI	300	89.3	91.3
Screenees			
All centers	911	84.0	89.0*
Detroit, MI	309	67.7	76.0
Pittsburgh, PA	300	89.3	95.3
Marshfield, WI	302	95.3	96.0

* $p \leq 0.05$ for difference between food frequency questionnaires.

TABLE 2. Mean percentages of missing or uninterpretable responses in a substudy of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, by type of food frequency questionnaire, 1998

	% of missing or uninterpretable responses					
	Questions on frequency of intake		Questions on portion size		Questions on use of dietary supplements	
	Diet History Questionnaire	PLCO food frequency questionnaire	Diet History Questionnaire	PLCO food frequency questionnaire	Diet History Questionnaire	PLCO food frequency questionnaire
Controls						
All centers	1.4	1.1	1.7	4.5***	6.0	14.7***
Detroit, MI	1.4	1.7	1.8	7.5***	7.2	19.5***
Pittsburgh, PA	1.4	0.8	2.0	4.0	7.0	9.3
Marshfield, WI	1.4	1.0	1.3	2.4**	4.1	15.3***
Screenees						
All centers	1.7***	0.6	2.0	2.8	5.4	6.4
Detroit, MI	2.3*	1.2	2.0	6.2*	6.6	13.9**
Pittsburgh, PA	1.1**	0.4	1.4	1.2	5.0***	0.6
Marshfield, WI	2.0***	0.4	2.5	2.0	4.8	7.0

* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$ (difference between food frequency questionnaires).

Table 2 shows the mean percentage of missing/uninterpretable responses by category of question for each of the two questionnaires, for both controls and screenees. For questions on frequency of intake among controls, the mean proportion of missing/uninterpretable responses was less than 2 percent overall, and the proportions were not significantly different by questionnaire type. Among screenees, use of the DHQ resulted in 1.7 percent of frequency-of-intake responses being missing/uninterpretable as compared with only 0.6 percent for the PLCO FFQ ($p \leq 0.001$). For questions on portion size (only considered when frequency of intake was recorded), the DHQ performed better than the PLCO FFQ ($p \leq 0.001$) among controls, with only 1.7 percent having missing/uninterpretable responses versus 4.5 percent (table 2). Similar though nonsignificant findings were seen among screenees. For dietary supplement questions, the mean percentage of missing/uninterpretable responses was 9 percentage points lower on the DHQ than on the PLCO FFQ among control participants ($p \leq 0.001$), with a similar but nonsignificant pattern being seen among screenees.

Table 3 shows responses to the question answered by control participants regarding the level of difficulty of each FFQ. Across all centers, participants seemed more likely to view the DHQ as "very easy" to complete than the PLCO FFQ, and this difference was significant for all centers combined. In addition, most control participants (approximately 90 percent) reported that they had needed no help to complete either questionnaire (data not shown).

DISCUSSION

These data show that the DHQ, a 36-page FFQ designed to be cognitively easier for respondents to complete, elicited a rate of response to an unannounced mailing among controls that was as high as that for the PLCO FFQ, which was less than half its length and required approximately 30 fewer minutes to complete. Given that the PLCO FFQ is similar to most FFQs currently being used in epidemiologic research, one could infer from these results that questionnaire length does not influence response rates for FFQs. However, given that the DHQ was developed with specific attention to cog-

TABLE 3. Responses of control participants (%) to questions on level of difficulty in a substudy of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, by type of food frequency questionnaire, 1998

	Difficulty of questionnaire					
	Very easy		Somewhat easy		Somewhat difficult or very difficult	
	Diet History Questionnaire	PLCO food frequency questionnaire	Diet History Questionnaire	PLCO food frequency questionnaire	Diet History Questionnaire	PLCO food frequency questionnaire
All centers*	58.0	50.1	37.0	40.1	5.0	9.8
Detroit, MI	65.4	52.2	27.9	37.8	6.7	9.9
Pittsburgh, PA	58.5	57.5	37.4	33.6	4.1	8.9
Marshfield, WI	51.5	42.2	43.8	47.4	4.6	10.4

* $p \leq 0.05$ for difference between food frequency questionnaires.

nitive ease, another possible explanation is that respondents are responsive to clarity even when a questionnaire is longer. The respondents' perceptions that the DHQ was easier to complete are consistent with this explanation. Other research has shown length to be an important but not necessarily determining factor in questionnaire response rates (3, 6, 7).

Among screenees, response rates between the two instruments were not as comparable. In this case, the study assessed response rates not only between instruments but also for a second administration of a detailed FFQ 3 years into the trial. For these participants, the DHQ was an added burden that always occurred after the PLCO FFQ administration. Furthermore, while the protocol for administering the DHQ at each center specified that the distribution and follow-up procedures be identical to those for the PLCO FFQ, it may have been difficult to replicate the same circumstances 3 years later, especially with respect to the amount of editing and data retrieval that occurred after questionnaire completion. Nevertheless, the findings for screenees, who had already exhibited high response rates for the PLCO FFQ, were encouraging and indicated a reasonable willingness to complete a second, long dietary instrument. Since baseline data collection, no other extensive questionnaires have been administered to PLCO participants. Therefore, there are no data with which to compare our declining response rates.

It is notable that the response rate for controls in this study was high in comparison with two recent large cohort studies that used mailings without prior notification—one in a multiethnic cohort with 2–3 follow-up mailings (response rates were 18.6–51.3 percent by gender/ethnic group) (11) and one among members of the American Association of Retired Persons with a single FFQ mailing (response rate = 17.6 percent) (A. Schatzkin, National Cancer Institute, personal communication, 2000). In the current study, the FFQs were mailed to a group of preselected individuals who had already agreed to participate in a screening trial. Therefore, response rates between this study and other large cohort studies in which FFQs are sent with little or no introduction and/or follow-up are not comparable. If the DHQ were mailed without prior notification, it is likely, given the response rates reported in recent cohort studies, that the response rates would be lower than those presented here. Introductory letters, personal contact, and multiple follow-ups tend to optimize response rates for any instrument and should be employed whenever possible (1, 2).

With respect to data quality, it is notable that the DHQ, while more than double the length of the PLCO FFQ, showed nearly the same small proportion of missing/uninterpretable responses for frequency-of-intake questions—data considered to be most critical for FFQs. For questions on portion size, the DHQ performed better, with a significantly smaller percentage of missing/uninterpretable responses among the controls, even though there were about 80 percent more portion size questions on the DHQ. This observation is important, because the grid format was deliberately abandoned in the DHQ partly to address the problems respondents have in answering two questions (frequency and portion size) on the same line (as is done in the PLCO FFQ). Although there is

disagreement among nutritional epidemiologists about the need to ask about portion size on FFQs (12), the fact remains that if the question is asked, obtaining answers is more likely with a nongrid format.

With respect to questions on the use of vitamin and mineral supplements, neither FFQ performed optimally, although the DHQ performed statistically significantly better, with 6.0 percent and 14.7 percent of responses being missing/uninterpretable for the DHQ and the PLCO FFQ, respectively. Both questionnaires could be improved in this area. Data on intake of vitamin and mineral supplements are notoriously difficult to collect (13). In both of these FFQs, the supplement questions appear on 2–3 pages at the end of the instrument; in the PLCO FFQ, the supplement questions are printed in a large grid which spans two pages horizontally. It appears that this format was difficult or confusing for many PLCO FFQ respondents. A general problem for both FFQs is that a number of respondents consume few, if any, of many of the supplements listed, and they tend to skip specifying "no" for supplements not taken. More thought and improvement in this area is necessary.

These data indicate the extent to which participants responded to questions on these specific FFQs. They do not provide any information about which questionnaire is more reproducible or valid with respect to estimating dietary intake. Clearly, one would not use a new, longer, and potentially more expensive questionnaire unless there were compelling reasons to do so. Data collection was recently completed in a validation study of the DHQ, comparing it with two other widely used FFQs and using 24-hour dietary recalls as reference data to address these issues. In the current study, we were simply assessing the practical feasibility of using the DHQ in the PLCO Trial by examining response rates and data quality, which are important factors in the performance of any instrument.

This study was designed to evaluate the feasibility of administering the recently developed DHQ to controls and screenees in the PLCO Trial. For both controls and screenees, the results showed that the DHQ provides reasonable data completeness. Among screenees only, participation rates for the DHQ, administered 3 years after the PLCO FFQ, were lower, indicating that response rates decline somewhat when a second comprehensive FFQ is administered to the same respondents. For controls, however, participation rates for a single first administration of the DHQ were excellent. This suggests that a shorter FFQ is not always better and that clarity and ease of administration may compensate for questionnaire length.

REFERENCES

1. Dillman DA. Mail and telephone surveys: the total design method. New York, NY: John Wiley and Sons, Inc, 1978.
2. Dillman DA. The design and administration of mail surveys. *Annu Rev Sociol* 1991;17:225–49.
3. Eaker S, Bergstrom R, Bergstrom A, et al. Response rate to mailed epidemiologic questionnaires: a population-based randomized trial of variations in design and mailing routines. *Am J Epidemiol* 1998;147:74–82.

4. Morris MC, Colditz GA, Evans DA. Response to a mail nutritional survey in an older bi-racial community population. *Ann Epidemiol* 1998;8:342-6.
5. Johansson L, Solvoll K, Opdahl S, et al. Response rates with different distribution methods and reward, and reproducibility of a quantitative food frequency questionnaire. *Eur J Clin Nutr* 1997;51:346-53.
6. Kristal AR, Glanz K, Feng Z, et al. Does using a short dietary questionnaire instead of a food frequency improve response rates to a health assessment survey? *J Nutr Educ* 1994;26:224-7.
7. Kuskowska-Wolk A, Holte S, Ohlander E-M, et al. Effects of different designs and extension of a food frequency questionnaire on response rate, completeness of data and food frequency responses. *Int J Epidemiol* 1992;21:1144-50.
8. Caan B, Hiatt RA, Owen A. Mailed dietary surveys: response rates, error rates and the effect of omitted food items on nutrient values. *Epidemiology* 1991;2:430-6.
9. Prorok PC, Andriole GL, Bresalier RS, et al. Design of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. *Controlled Clin Trials* 2000;21(suppl):273s-309s.
10. Subar AF, Thompson FE, Smith AF, et al. Improving food frequency questionnaires: a qualitative approach using cognitive interviewing. *J Am Diet Assoc* 1995;95:781-8.
11. Kolonel LN, Henderson BE, Hankin JH, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am J Epidemiol* 2000;151:346-57.
12. Willett WC. *Nutritional epidemiology*. New York, NY: Oxford University Press, 1998.
13. Patterson RE, Kristal AR, Levy L, et al. Validity of methods used to assess vitamin and mineral supplement use. *Am J Epidemiol* 1998;148:643-9.



PRACTICE OF EPIDEMIOLOGY

Is Shorter Always Better? Relative Importance of Questionnaire Length and Cognitive Ease on Response Rates and Data Quality for Two Dietary Questionnaires

Amy F. Subar,¹ Regina G. Ziegler,² Frances E. Thompson,¹ Christine Cole Johnson,³ Joel L. Weissfeld,⁴ Douglas Reding,⁵ Katherine H. Kavounis,⁶ and Richard B. Hayes⁷ for the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial Investigators

In this study, the authors sought to determine the effects of length and clarity on response rates and data quality for two food frequency questionnaires (FFQs): the newly developed 36-page Diet History Questionnaire (DHQ), designed to be cognitively easier for respondents, and a 16-page FFQ developed earlier for the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. The PLCO Trial is a 23-year randomized controlled clinical trial begun in 1992. The sample for this substudy, which was conducted from January to April of 1998, consisted of 900 control and 450 screened PLCO participants aged 55–74 years. Controls received either the DHQ or the PLCO FFQ by mail. Screenees, who had previously completed the PLCO FFQ at baseline, were administered the DHQ. Among controls, the response rate for both FFQs was 82%. Average amounts of time needed by controls to complete the DHQ and the PLCO FFQ were 68 minutes and 39 minutes, respectively. Percentages of missing or uninterpretable responses were similar between instruments for questions on frequency of intake but were approximately 3 and 9 percentage points lower ($p \leq 0.001$) in the DHQ for questions on portion size and use of vitamin/mineral supplements, respectively. Among screenees, response rates for the DHQ and the PLCO FFQ were 84% and 89%, respectively, and analyses of questions on portion size and supplement use showed few differences. These data indicated that the shorter FFQ was not better from the perspective of response rate and data quality, and that clarity and ease of administration may compensate for questionnaire length. *Am J Epidemiol* 2001;153:404–9.

diet; diet surveys; epidemiologic methods; food habits; nutrition assessment; nutrition surveys; questionnaires

In epidemiologic and public health research, cooperation in completing questionnaires is elicited in a variety of ways. Depending on the research questions, study design, funding, and practical considerations, questionnaires can be

interviewer- or self-administered, long or short, easy or complex. Efforts to maximize response rates generally include some combination of introductions, reminders, and follow-ups through in-person, telephone, or mail contacts. Incentives such as payment, gifts, or medical information may also enhance final response rates. Response rates, in turn, are important measures of study success, validity, and generalizability.

Following telephone data collection methods, mail surveys are the most common method of obtaining questionnaire data (1, 2). They are economical and practical in many research situations and are common in large-scale epidemiologic research. A review of past research indicates that for mailed surveys, follow-up is the most important factor in increasing response rates, followed by prior notice, financial incentives, special mailing, a positive perception of the survey sponsors, provision of a postage-paid return envelope, and personalization (2). It is less clear how questionnaire length affects response rates.

Nutritional epidemiologic studies frequently incorporate mailed, self-administered questionnaires to collect dietary data, particularly in large prospective studies or intervention trials in which participants are spread over a wide geo-

Received for publication January 31, 2000, and accepted for publication May 26, 2000.

Abbreviations: DHQ, Diet History Questionnaire; FFQ, food frequency questionnaire; PLCO, Prostate, Lung, Colorectal, and Ovarian [Cancer Screening Trial].

¹Risk Factor Monitoring and Methods Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD.

²Epidemiology and Biostatistics Program, Division of Cancer Etiology and Genetics, National Cancer Institute, Bethesda, MD.

³Josephine Ford Cancer Center, Detroit, MI.

⁴Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA.

⁵Marshfield Medical Research and Education Foundation, Marshfield, WI.

⁶Westat, Incorporated, Rockville, MD.

⁷Environmental Epidemiology Branch, Division of Cancer Etiology and Genetics, National Cancer Institute, Bethesda, MD.

Reprint requests to Dr. Amy F. Subar, National Cancer Institute, Executive Plaza North, Room 4005, 6130 Executive Blvd., MSC 7344, Bethesda, MD 20892-7344 (e-mail: amy_subar@nih.gov).

graphic area. Many such studies rely on the use of machine-readable food frequency questionnaires (FFQs), which generally require a substantial time commitment (30–60 minutes) from participants. Several studies have addressed various methodological issues related to response rates and data quality specifically for mailed, self-administered FFQs (3–8). Similar to findings for mailed surveys in general, factors such as preliminary notification, personal contact, incentives, and follow-up contacts increase response rates for FFQs (3–5), but findings regarding the effect of questionnaire length and design on response rates or data quality have been inconsistent (3, 6–8).

Although data conflict regarding whether questionnaire length actually decreases response rates, concerns about this and increased respondent burden can preclude or limit detailed dietary assessment in nutritional epidemiologic studies. In this analysis, we examined response rates and data quality between two dietary questionnaires: a new 36-page FFQ designed to be cognitively easier for respondents and a 16-page FFQ. The longer questionnaire incorporated many elements of the “total design method” for mail surveys developed by Dillman (1, 2), the purpose of which is to maximize response rates and data quality. The method attends in detail to visible and cognitive aspects of the questionnaire development and survey implementation process, including the ordering of questions, graphic design, and content.

MATERIALS AND METHODS

Study design

The Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, begun in 1992, is a 23-year randomized controlled clinical trial of screening procedures for prostate, lung, colon, and ovarian cancer involving 148,000 men and women aged 55–74 years. The 10 screening centers for the trial are the Georgetown University Medical Center (Washington, DC), the Henry Ford Health System (Detroit, Michigan), the Marshfield Medical Education and Research Foundation (Marshfield, Wisconsin), the Pacific Health Research Institute (Honolulu, Hawaii), the University of Alabama (Birmingham, Alabama), the University of Colorado (Denver, Colorado), the University of Minnesota (Minneapolis, Minnesota), the University of Pittsburgh (Pittsburgh, Pennsylvania), the University of Utah (Salt Lake City, Utah), and Washington University (St. Louis, Missouri). Participants have been randomized into either the screened arm or the control arm of the trial. Details regarding the study’s design and purpose are reported elsewhere (9).

Additional etiologic studies of cancer were developed in both the screening and control components of the trial. Diet was one of a number of exposures to be assessed. All participants randomized into the screening arm of the trial received a PLCO FFQ (described below) at baseline either by mail or in person (depending on the practice specific to each site).

The current substudy, conducted from January to April of 1998, resulted from an effort to determine the practical utility of adding a recently developed FFQ, the Diet

History Questionnaire (DHQ) (described below), to both the control and the screening arms of the PLCO Trial. The new instrument is more detailed and cognitively based than the original PLCO FFQ, and therefore we thought that it might be advantageous to include it in the study. Testing of the willingness of both screenees and controls to complete the instrument was considered essential in deciding whether or not to add the DHQ to the study protocol. Three screening centers were selected to participate in this analysis: the Henry Ford Health System, the University of Pittsburgh, and the Marshfield Medical Education and Research Foundation. At each center, 300 control participants (a total of 900) were randomly selected; half were randomized to receive, by mail, the DHQ and half to receive the PLCO FFQ. The mailing, conducted by each center, was accompanied by a cover letter and a postage-paid return envelope. For participants who did not return their questionnaires within 3 weeks, up to five telephone calls were made by staff at each center (at different times and days of the week) in an effort to make a single follow-up contact. A second questionnaire was provided to participants who reported losing or misplacing the first questionnaire. In addition to either the DHQ or the PLCO FFQ, participants were given a one-page questionnaire on which they recorded the amount of time they had needed to complete the questionnaire. The participants were asked to answer two qualitative questions: 1) “Which of the following categories best describes how much help you needed in answering how often you eat foods or your usual portion size?” (response categories: no help needed, little help needed, moderate help needed, much help needed) and 2) “How easy or difficult was the dietary questionnaire to fill out?” (response categories: very easy, somewhat easy, somewhat difficult, very difficult).

In the screening arm of the study, 150 participants at each of the three screening centers (a total of 450), who had previously received a PLCO FFQ at baseline as part of the trial protocol, were selected sequentially as their year 3 clinic visits occurred to receive the DHQ, to be completed either on-site or at home (with a postage-paid return envelope). Whichever procedures each screening center had previously used for administration and follow-up of the PLCO FFQ at baseline were adhered to for the DHQ administration in year 3. Screened participants were not asked the questions about length and difficulty of the dietary instrument, because there were no comparable data from the administration of the PLCO FFQ at baseline.

Food frequency questionnaires

In 1994, investigators at the National Cancer Institute began to develop a new, self-administered, machine-readable FFQ in the hope of improving the FFQs currently being used in epidemiologic research. The intent was to improve both the cognitive and database aspects of FFQs. To this end, cognitive testing was conducted (10) which resulted in the creation of the DHQ, a 36-page FFQ that was improved with respect to content, order, wording, and

layout. The DHQ asked about frequency of intake for 114 individual food items; for 109 of the items, the DHQ asked about portion size by providing a choice of three portion size ranges. For 47 of these food items, 1–7 additional embedded questions asked about related aspects of consumption, such as seasonal intake, food type (e.g., low-fat, lean, diet, caffeine-free), fat additions, or fat type. The DHQ also asked 1) five questions about the proportion of the time the respondent used various types of margarine, mayonnaise, sour cream, cream cheese, and salad dressing; 2) four summary questions; and 3) nine questions on the use of vitamin and mineral supplements, six of which queried about frequency of intake and duration and five of which queried about usual dose. A copy of the DHQ can be found at <http://www-dccps.ims.nci.nih.gov/ARP/DHQ>.

In 1993, prior to completion of the DHQ in 1998, a shorter FFQ that incorporated some of the DHQ innovations was developed for the PLCO Trial. This PLCO FFQ is a 16-page machine-readable booklet with the appearance of a standard Block/National Cancer Institute or Willett grid-format FFQ. It was developed to characterize total diet while minimizing participant burden. The PLCO FFQ contains questions on 137 individual food items, 60 of which include a query about usual portion size; 10 questions about meat cooking practices; and 14 questions about intake of vitamin/mineral supplements, 12 of which contain detailed queries regarding current use, dosage, duration, and use 2 and 5 years previously. Overall, the PLCO FFQ included many fewer items overall and many more items per page than the DHQ. A complete copy of the PLCO FFQ can be found at <http://dcp.nci.nih.gov/plco/diet>.

Analyses

Differences between the two questionnaires for each arm of the trial were assessed by comparing response rates, percentages of unusable responses, and, for the control arm, times needed to complete the questionnaire and answers to qualitative questions. Absolute numbers for questions on frequency of intake, portion size, and supplement use differed between the two FFQs; therefore, for those three types of questions, the mean percentages of questionnaire responses that were missing or uninterpretable (because of stray marks or double-responding found in the machine scanning) were compared using *t* tests among controls and paired *t* tests among screenees. We conducted χ^2 tests to determine differences in response rates and in answers to qualitative questions between the DHQ and PLCO FFQ. A *t* test was used to determine differences between DHQ and PLCO FFQ respondents in the amount of time taken to complete the instruments.

RESULTS

The study sample of controls and screenees combined was 51 percent female and 49 percent male. Fourteen percent of participants were aged 50–59 years, 54 percent were aged 60–69 years, and 32 percent were aged 70 years or more. Approximately 11 percent had less than a high school

education, 29 percent had graduated from high school, 33 percent had some post-high-school education, and 27 percent were college graduates or had postcollege education. Randomization of the controls to receive either the DHQ or the PLCO FFQ showed that the participants were evenly distributed by gender and were within 3 percentage points by age group.

Table 1 shows the percentages of DHQs and PLCO FFQs returned by controls and screenees. Among controls, the response rates for the two questionnaires for all centers combined were nearly identical (approximately 82 percent). Within centers, response rates for controls by FFQ type differed by 0.7–8.3 percentage points. None of the differences were statistically significant. In the screenee group, the percentage of participants who returned the DHQ was 5 percentage points lower than the PLCO FFQ return rate for these same participants 3 years earlier (84 percent vs. 89 percent) ($p \leq 0.05$). Furthermore, the DHQ return rate by center was consistently lower than the earlier return rate for the PLCO FFQ.

After initial response rates were determined, a decision was made to exclude respondents for whom 33 percent of frequency-of-intake responses were missing or uninterpretable (at or above the 99th percentile of the frequency distribution). Four of the DHQs and five of the PLCO FFQs returned by controls were excluded, as were five DHQs and no PLCO FFQs returned by screenees. Finally, because the analyses for screenees required within-subject comparisons, 21 persons were excluded because they had not successfully completed both the DHQ and the PLCO FFQ by the above exclusion criterion. This left 357 screenees for the final analyses.

The mean amounts of time taken by controls to complete the DHQ and the PLCO FFQ were 68 minutes and 39 minutes, respectively. The times varied little by center and were significantly different overall and for each center ($p \leq 0.0001$).

TABLE 1. Percentages of food frequency questionnaires returned by control and screened participants in a substudy of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, 1998

	% of questionnaires returned		
	No. of participants	Diet History Questionnaire	PLCO food frequency questionnaire
Controls			
All centers	900	81.9	81.6
Detroit, MI	300	72.3	75.0
Pittsburgh, PA	300	84.0	78.7
Marshfield, WI	300	89.3	91.3
Screenees			
All centers	911	84.0	89.0*
Detroit, MI	309	67.7	76.0
Pittsburgh, PA	300	89.3	95.3
Marshfield, WI	302	95.3	96.0

* $p \leq 0.05$ for difference between food frequency questionnaires.

TABLE 2. Mean percentages of missing or uninterpretable responses in a substudy of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, by type of food frequency questionnaire, 1998

	% of missing or uninterpretable responses					
	Questions on frequency of intake		Questions on portion size		Questions on use of dietary supplements	
	Diet History Questionnaire	PLCO food frequency questionnaire	Diet History Questionnaire	PLCO food frequency questionnaire	Diet History Questionnaire	PLCO food frequency questionnaire
Controls						
All centers	1.4	1.1	1.7	4.5***	6.0	14.7***
Detroit, MI	1.4	1.7	1.8	7.5***	7.2	19.5***
Pittsburgh, PA	1.4	0.8	2.0	4.0	7.0	9.3
Marshfield, WI	1.4	1.0	1.3	2.4**	4.1	15.3***
Screenees						
All centers	1.7***	0.6	2.0	2.8	5.4	6.4
Detroit, MI	2.3*	1.2	2.0	6.2*	6.6	13.9**
Pittsburgh, PA	1.1**	0.4	1.4	1.2	5.0***	0.6
Marshfield, WI	2.0***	0.4	2.5	2.0	4.8	7.0

* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$ (difference between food frequency questionnaires).

Table 2 shows the mean percentage of missing/uninterpretable responses by category of question for each of the two questionnaires, for both controls and screenees. For questions on frequency of intake among controls, the mean proportion of missing/uninterpretable responses was less than 2 percent overall, and the proportions were not significantly different by questionnaire type. Among screenees, use of the DHQ resulted in 1.7 percent of frequency-of-intake responses being missing/uninterpretable as compared with only 0.6 percent for the PLCO FFQ ($p \leq 0.001$). For questions on portion size (only considered when frequency of intake was recorded), the DHQ performed better than the PLCO FFQ ($p \leq 0.001$) among controls, with only 1.7 percent having missing/uninterpretable responses versus 4.5 percent (table 2). Similar though nonsignificant findings were seen among screenees. For dietary supplement questions, the mean percentage of missing/uninterpretable responses was 9 percentage points lower on the DHQ than on the PLCO FFQ among control participants ($p \leq 0.001$), with a similar but nonsignificant pattern being seen among screenees.

Table 3 shows responses to the question answered by control participants regarding the level of difficulty of each FFQ. Across all centers, participants seemed more likely to view the DHQ as "very easy" to complete than the PLCO FFQ, and this difference was significant for all centers combined. In addition, most control participants (approximately 90 percent) reported that they had needed no help to complete either questionnaire (data not shown).

DISCUSSION

These data show that the DHQ, a 36-page FFQ designed to be cognitively easier for respondents to complete, elicited a rate of response to an unannounced mailing among controls that was as high as that for the PLCO FFQ, which was less than half its length and required approximately 30 fewer minutes to complete. Given that the PLCO FFQ is similar to most FFQs currently being used in epidemiologic research, one could infer from these results that questionnaire length does not influence response rates for FFQs. However, given that the DHQ was developed with specific attention to cog-

TABLE 3. Responses of control participants (%) to questions on level of difficulty in a substudy of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, by type of food frequency questionnaire, 1998

	Difficulty of questionnaire					
	Very easy		Somewhat easy		Somewhat difficult or very difficult	
	Diet History Questionnaire	PLCO food frequency questionnaire	Diet History Questionnaire	PLCO food frequency questionnaire	Diet History Questionnaire	PLCO food frequency questionnaire
All centers*	58.0	50.1	37.0	40.1	5.0	9.8
Detroit, MI	65.4	52.2	27.9	37.8	6.7	9.9
Pittsburgh, PA	58.5	57.5	37.4	33.6	4.1	8.9
Marshfield, WI	51.5	42.2	43.8	47.4	4.6	10.4

* $p \leq 0.05$ for difference between food frequency questionnaires.

nitive ease, another possible explanation is that respondents are responsive to clarity even when a questionnaire is longer. The respondents' perceptions that the DHQ was easier to complete are consistent with this explanation. Other research has shown length to be an important but not necessarily determining factor in questionnaire response rates (3, 6, 7).

Among screenees, response rates between the two instruments were not as comparable. In this case, the study assessed response rates not only between instruments but also for a second administration of a detailed FFQ 3 years into the trial. For these participants, the DHQ was an added burden that always occurred after the PLCO FFQ administration. Furthermore, while the protocol for administering the DHQ at each center specified that the distribution and follow-up procedures be identical to those for the PLCO FFQ, it may have been difficult to replicate the same circumstances 3 years later, especially with respect to the amount of editing and data retrieval that occurred after questionnaire completion. Nevertheless, the findings for screenees, who had already exhibited high response rates for the PLCO FFQ, were encouraging and indicated a reasonable willingness to complete a second, long dietary instrument. Since baseline data collection, no other extensive questionnaires have been administered to PLCO participants. Therefore, there are no data with which to compare our declining response rates.

It is notable that the response rate for controls in this study was high in comparison with two recent large cohort studies that used mailings without prior notification—one in a multiethnic cohort with 2–3 follow-up mailings (response rates were 18.6–51.3 percent by gender/ethnic group) (11) and one among members of the American Association of Retired Persons with a single FFQ mailing (response rate = 17.6 percent) (A. Schatzkin, National Cancer Institute, personal communication, 2000). In the current study, the FFQs were mailed to a group of preselected individuals who had already agreed to participate in a screening trial. Therefore, response rates between this study and other large cohort studies in which FFQs are sent with little or no introduction and/or follow-up are not comparable. If the DHQ were mailed without prior notification, it is likely, given the response rates reported in recent cohort studies, that the response rates would be lower than those presented here. Introductory letters, personal contact, and multiple follow-ups tend to optimize response rates for any instrument and should be employed whenever possible (1, 2).

With respect to data quality, it is notable that the DHQ, while more than double the length of the PLCO FFQ, showed nearly the same small proportion of missing/uninterpretable responses for frequency-of-intake questions—data considered to be most critical for FFQs. For questions on portion size, the DHQ performed better, with a significantly smaller percentage of missing/uninterpretable responses among the controls, even though there were about 80 percent more portion size questions on the DHQ. This observation is important, because the grid format was deliberately abandoned in the DHQ partly to address the problems respondents have in answering two questions (frequency and portion size) on the same line (as is done in the PLCO FFQ). Although there is

disagreement among nutritional epidemiologists about the need to ask about portion size on FFQs (12), the fact remains that if the question is asked, obtaining answers is more likely with a nongrid format.

With respect to questions on the use of vitamin and mineral supplements, neither FFQ performed optimally, although the DHQ performed statistically significantly better, with 6.0 percent and 14.7 percent of responses being missing/uninterpretable for the DHQ and the PLCO FFQ, respectively. Both questionnaires could be improved in this area. Data on intake of vitamin and mineral supplements are notoriously difficult to collect (13). In both of these FFQs, the supplement questions appear on 2–3 pages at the end of the instrument; in the PLCO FFQ, the supplement questions are printed in a large grid which spans two pages horizontally. It appears that this format was difficult or confusing for many PLCO FFQ respondents. A general problem for both FFQs is that a number of respondents consume few, if any, of many of the supplements listed, and they tend to skip specifying "no" for supplements not taken. More thought and improvement in this area is necessary.

These data indicate the extent to which participants responded to questions on these specific FFQs. They do not provide any information about which questionnaire is more reproducible or valid with respect to estimating dietary intake. Clearly, one would not use a new, longer, and potentially more expensive questionnaire unless there were compelling reasons to do so. Data collection was recently completed in a validation study of the DHQ, comparing it with two other widely used FFQs and using 24-hour dietary recalls as reference data to address these issues. In the current study, we were simply assessing the practical feasibility of using the DHQ in the PLCO Trial by examining response rates and data quality, which are important factors in the performance of any instrument.

This study was designed to evaluate the feasibility of administering the recently developed DHQ to controls and screenees in the PLCO Trial. For both controls and screenees, the results showed that the DHQ provides reasonable data completeness. Among screenees only, participation rates for the DHQ, administered 3 years after the PLCO FFQ, were lower, indicating that response rates decline somewhat when a second comprehensive FFQ is administered to the same respondents. For controls, however, participation rates for a single first administration of the DHQ were excellent. This suggests that a shorter FFQ is not always better and that clarity and ease of administration may compensate for questionnaire length.

REFERENCES

1. Dillman DA. Mail and telephone surveys: the total design method. New York, NY: John Wiley and Sons, Inc, 1978.
2. Dillman DA. The design and administration of mail surveys. *Annu Rev Sociol* 1991;17:225–49.
3. Eaker S, Bergstrom R, Bergstrom A, et al. Response rate to mailed epidemiologic questionnaires: a population-based randomized trial of variations in design and mailing routines. *Am J Epidemiol* 1998;147:74–82.

4. Morris MC, Colditz GA, Evans DA. Response to a mail nutritional survey in an older bi-racial community population. *Ann Epidemiol* 1998;8:342-6.
5. Johansson L, Solvoll K, Opdahl S, et al. Response rates with different distribution methods and reward, and reproducibility of a quantitative food frequency questionnaire. *Eur J Clin Nutr* 1997;51:346-53.
6. Kristal AR, Glanz K, Feng Z, et al. Does using a short dietary questionnaire instead of a food frequency improve response rates to a health assessment survey? *J Nutr Educ* 1994;26:224-7.
7. Kuskowska-Wolk A, Holte S, Ohlander E-M, et al. Effects of different designs and extension of a food frequency questionnaire on response rate, completeness of data and food frequency responses. *Int J Epidemiol* 1992;21:1144-50.
8. Caan B, Hiatt RA, Owen A. Mailed dietary surveys: response rates, error rates and the effect of omitted food items on nutrient values. *Epidemiology* 1991;2:430-6.
9. Prorok PC, Andriole GL, Bresalier RS, et al. Design of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. *Controlled Clin Trials* 2000;21(suppl): 273s-309s.
10. Subar AF, Thompson FE, Smith AF, et al. Improving food frequency questionnaires: a qualitative approach using cognitive interviewing. *J Am Diet Assoc* 1995;95:781-8.
11. Kolonel LN, Henderson BE, Hankin JH, et al. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am J Epidemiol* 2000;151:346-57.
12. Willett WC. *Nutritional epidemiology*. New York, NY: Oxford University Press, 1998.
13. Patterson RE, Kristal AR, Levy L, et al. Validity of methods used to assess vitamin and mineral supplement use. *Am J Epidemiol* 1998;148:643-9.