Gender Differences in Social Desirability and Social Approval Bias in Dietary Self-report

James R. Hebert,[1] Yunsheng Ma,[1] Lynn Clemow,[1] Ira S. Ockene,[1] Gordon Saperia,[2] Edward J. Stanek III,[3] Philip A. Merriam,[1] and Judith K. Ockene[1]

Social desirability (the tendency to respond in such a way as to avoid criticism) and social approval (the tendency to seek praise) are two prominent response set biases evident in answers on structured questionnaires. These biases were tested by comparing nutrient intakes as estimated from a single 24-hour diet recall interview (24HR) and a 7-day dietary recall (7DDR). Data were collected as part of the Worcester Area Trial for Counseling in Hyperlipidemia, a randomized, physician-delivered nutrition intervention trial for hypercholesterolemic patients conducted in Worcester, Massachusetts, from 1991 to 1995. Of the 1,278 total study subjects, 759 had complete data for analysis. Men overestimated their fat and energy intakes on the 7DDR as compared with the 24HR according to social approval: One unit increase in the social approval score was associated with an overestimate of 21.5 kcal/day in total energy intake and 1.2 g/day in total fat intake. Women, however, underestimated their dietary intakes on the 7DDR relative to the 24HR according to social desirability: One unit increase in the social desirability score was associated with an underestimate of 19.2 kcal/day in energy intake and 0.8 g/day in total fat. The results from the present study indicate that social desirability and social approval biases appear to vary by gender. Such biases may lead to misclassification of dietary exposure estimates resulting in a distortion in the perceived relation between health-related outcomes and exposure to specific foods or nutrients. Because these biases may vary according to the perceived demands of research subjects, it is important that they be assessed in a variety of potential research study populations. *Am J Epidemiol* 1997;146:1046–55.

bias (epidemiology); diet; food habits; nutrition assessment; personality assessment; questionnaires; social desirability

Dietary self-report measures are a mainstay of epidemiologic research on diet and disease relations. However, the validity of these measures has received relatively little attention. The reporting of dietary behavior is a complex process, including cognitive, perceptual, and emotional aspects. Studies on the accuracy of dietary measures have focused primarily on the processes of memory (e.g., the strong influence of recent intake on summary self-report) (1–4) and perception (e.g., validity and reliability in quantity estimation) (5–7). Other potential sources of error in dietary self-report generally have gone unrecognized or have been assumed to occur randomly across populations. The lack of attention to examining biases in dietary self-report is interesting when contrasted to other disciplines in which measurement also relies on self-reports of behavior. In psychological research, for example, the possibility of a self-report measure being open to many complex sources of bias is assumed to exist unless proven otherwise (8, 9). One essential response set to check on any self-report measure is social desirability, which has been ignored in dietary self-report until very recently.

Social desirability and social approval

Social desirability is the defensive tendency of individuals to respond in a manner consistent with societal norms or beliefs. Originally investigated by Edwards (10, 11), social desirability was seen as an "attribute of test items." Certain test items, because of their content or phrasing, provided choices that were

generally perceived as being "much better than others." Social desirability eventually evolved to be seen as a characteristic of the subject, or at least the person's test behavior (12). Individuals scoring high on social desirability are far more acquiescent to perceived situational demands in experimental settings (12). The 33 true-false questions on the Marlowe-Crowne Social Desirability scale (MCSD) present the respondent with extreme positions on topics that tend to be very charged (e.g., "I never resent being asked to return a favor"). By contrast, social approval is the tendency for an individual to seek a positive response in testing situations and therefore is less focused on defensiveness (13). The 20 questions on the Martin-Larsen Approval Motivation (MLAM) scale are less extreme and provide five-level Likert scale responses rather than the simple dichotomies of the MCSD (e.g., "I am willing to argue only if I know that my friends will back me up"). Both social desirability and approval scales have been used widely in psychological research including studies aiming to refine psychological test instruments or to provide a means for adjusting test scores. Such adjustments are commonly ignored in epidemiologic research employing self-assessments even though biases may be large for self-reports of certain behaviors such as alcohol consumption, smoking, sexual behavior, and diet.

The potential for social desirability and social approval response bias regarding food intake is great because people have strong emotional associations with food (14) and certain foods are widely recognized as either "good for you" or "bad for you" (15). Consequently, it is now widely held that some foods, such as fruit and vegetables, are healthy, whereas other foods, such as those rich in fat, are not healthy.

Biases due to social desirability and social approval often are evident in situations that can be perceived as a test and in responding to questions for which a particular response is more socially acceptable than another (10–13). Dietary self-assessment instruments such as the food frequency questionnaire (FFQ) may be viewed as "tests" by study subjects. Demographic groups often produce estimates of nutrient intakes derived from the FFQ that, in crude terms, are biased in a manner consistent with societal norms. Pregnant women tend to overestimate food intake in keeping with antenatal dietary advice (16) whereas a general female population tends to underestimate (17). Despite their practical appeal for establishing long-term habitual intake (18), the potential for bias in these methods may be greater than for shorter-period assessments such as food diaries or 24-hour diet recall interviews (24HR).

For structured questionnaires aiming to establish long-term dietary patterns, the participant must invoke habitual memory of dietary intake over the relevant period and do the arithmetic necessary to estimate average exposure over well-defined periods. This is a difficult cognitive task. In general, it has been shown that recall of past diet may be subject to a number of biases (19, 20). It also has been shown that specific dietary exposures are replaced by more generic memory as duration of the inquiry period increases (1–3, 21, 22). For periods of more than 2 weeks, generic memory predominates over episodic memory. Work by Smith, Jobe, and Mingay (2) indicates that the contextual frame in which questions are asked influences responses to questions, especially regarding food frequency. The first two stages in the cognitive process, question comprehension and information retrieval, might vary predictably by factors such as educational attainment and age but probably are less affected by factors such as social norms (20). However, nonexplicit messages (e.g., factors related to social norms) may condition generic memory of food encounters, essentially altering the third of the four main stages in the cognitive process, estimation/judgment (22). The fourth stage, response formulation, also may be modified by the individual perception of the consequences of an "unacceptable" answer (1, 22). A report on long-term average intake may be more prone to such biases because it is much more a statement of one's habitual dietary behavior than is reported intake of specific foods either while they are being eaten or during a short period in the very recent past. By providing the necessary list of foods, which is avoided entirely in diet recalls or records, there is an effective prompting of responses to specific foods that, for a variety of cultural and psychological reasons, are easily recognizable as good (or desirable) or bad (undesirable).

Gender differences

Women's responses to dietary questionnaires may be influenced more powerfully by social desirability than those of men, and women are generally higher in social desirability as a trait than are men (23). Social desirability is more likely to be a biasing factor in situations in which individuals experience some conflict between their true preferences and the socially desirable answer. Conflict and guilt around consumption of these foods is expressed early in childhood, especially by girls (24).

It must be emphasized that items or issues that may be neutral for one population may be important for another (8). Well-documented sex differences on attitudes about food might mediate social motivations to

alter responses. Recent figures suggest that between 33 and 40 percent of women currently are dieting—nearly twice the rate of men—and an additional 28 percent are trying to maintain a weight loss (25). Dieting for weight loss clearly increases awareness of food and increases the chances that the individual will adopt judgmental attitudes about departures from their eating plans (26). Dietary restraint, much more prevalent in women than in men, is a related concept that describes a state of chronic weight consciousness and a chronic pattern of attempting to exert self-control in eating (27). Consequently, women are more likely to feel guilty about eating (28). As one study conducted in a cafeteria setting has shown, individuals high in restraint underestimated intake even when the caloric values of the food choices were provided (29).

In a previous study (23), our group observed a large downward bias due to social desirability in comparing estimates of total energy and fat intake using a 7-day diet recall (7DDR) (similar in some ways to an FFQ) with those derived from multiple 24HRs. The bias was two to three times larger in women than in men. Because of the modest sample size of that study (41 subjects) and the unique volunteer aspect of its population, we were motivated to reexamine the potential roles of these biases in a larger and less health-conscious population. This paper reports results from the Worcester Area Trial for Counseling in Hyperlipidemia (WATCH). We quantified social desirability and social approval bias in the WATCH study by comparing nutrient intakes derived from the self-administered 7DDR and a telephone-administered 24HR. The usefulness of the 24HR as a comparison method is based on the fact that a short-term method depends nearly entirely on episodic memory (2) and is free from the constraint of prompting by food lists (23).

MATERIALS AND METHODS

The WATCH was a National Heart, Lung, and Blood Institute-supported randomized clinical trial testing the effectiveness of a physician-delivered nutrition intervention counseling program in reducing dietary fat and low density lipoprotein cholesterol levels in hyperlipidemic patients. A total of 1,278 patients in the upper quartile of the cholesterol distribution were recruited into the study (30).

Data collection

Demographic information. A wide variety of demographic data was collected at baseline including age, gender, years of education, race, and marital status.

Nondietary aspects of lifestyle and psychosocial data. At baseline and at 1 year, after the intervention was completed, data were collected on dietary preferences and motivations, self-efficacy, health beliefs, stages of change, and smoking status. After the intervention, the MCSD (12) and the MLAM (31) scales, instruments with high internal consistency and test-retest reliability (31), were administered to all study subjects. The MCSD scale consists of 33 true or false questions. One point is scored for each instance in which a true or false response matches on the scoring algorithm. The sum of all the matched responses is the social desirability score. The MLAM scale is a 20-question instrument requiring a five-point scale response. All questions are scored in the way they appear except questions 2, 12, 13, 16, and 19, which are reverse-scored. The sum of the individual response scores is the social approval score.

The MCSD and MLAM scales were mailed with a letter and a self-addressed stamped envelope to all WATCH study participants in March 1995. One month after the first mailing, another letter was sent as a reminder to the participants who had not responded to the first mailing. A third letter, mailed a month later and including a $5 incentive offer, was sent to participants who had not responded to the first two letters. One month later, a fourth letter was sent to nonresponders to the first three letters, offering a $10 incentive.

Nutritional assessment. All 24HRs were conducted by research dieticians using a computerized telephone survey method. The Nutrition Data System (32) was used for both interviewing and nutrient computation. A single 24HR per participant was administered at baseline and then 1 year later, each on randomly selected days. The 7DDR was developed to assess intake of more than 95 percent of dietary lipids (including saturated fat) and calories in our study population (33). It accounts for more than 85 percent of many other nutrients of public health interest. In certain respects, the 7DDR looks very much like an FFQ. It is a 118-food and 13-beverage item questionnaire, with worksheet, that asks participants to recall specific meals and snacks over the past week. In instances in which specific recall of a dietary encounter is not possible, subjects are requested to provide their typical or average intake, in the same way as subjects do on an FFQ. Thus, the 7DDR mixes, to a varying extent, information retrieved from episodic and generic memory. Nutrient scores derived from the 24HR and the 7DDR were computed from the same Nutrition Data System database.

In addition to dietary data information, height and weight data were collected at baseline, when study

participants were recruited during 1990–1993. Body mass index (BMI) was computed using the following formula: weight(kg)/height(m)^2.

Statistical methods

The purpose of this analysis was to investigate whether either social desirability or social approval is associated with biased estimates of fat and total energy intakes at baseline. Analyses were conducted using the SAS system (34, 35). Preliminary analyses evaluated univariate statistics and distributions for each variable to check for potential outliers. Pearson product moment correlations were used to screen for highly correlated variables. Demographic and nutrition variables were compared between responder (who completed the social approval and desirability scales) and nonresponders using t tests (for continuous variables) and chi-square tests (for categorical variables). These comparisons also were stratified by gender.

To assess the role of social desirability or social approval in biasing estimates of nutrient intake on the 7DDR versus the 24HR, we used a general linear model (PROC GLM) in SAS (34). Models were fit with the 7DDR-derived nutrient score as the dependent variable. The corresponding 24HR-derived nutrient score was fitted as an independent variable, as were the following: social desirability and social approval scores, information on whether patients received an incentive for completing the social desirability and social approval forms, interval (days between mailing of the forms and their return), and BMI. BMI and interval previously had been observed to be associated with differences in nutrient scores obtained from different dietary assessment methods. Because females and males appeared to provide qualitatively different answers both in previous analyses (23) and in analyses of these data (i.e., there was effect modification by gender), models were stratified by gender.

RESULTS

Of the 1,278 WATCH study participants, 759 completed the MCSD and MLAM forms and had dietary data necessary for these analyses. This represents a response rate of 81.3 percent among individuals with all dietary data necessary for analysis ($n = 934$) and an overall response rate of 59.4 percent. As shown in table 1, responders were more likely to be female, married, white, and unemployed and to have more nonstudy lipid measures performed, more education, and a higher baseline serum cholesterol concentration. Descriptive analyses of data from responders indicated

that all continuous variables were approximately normally distributed and otherwise adhered to assumptions of the linear regression models. Correlations between candidate independent variables were consistently low (i.e., $r \leq |0.25|$). Social desirability and social approval scores were uncorrelated ($r = -0.17$). As is typically the case with dietary data, dietary fat explained most of the variability in total energy intake (regression model $R^2 = 0.83$ for 7DDR data and $R^2 = 0.70$ for 24HR data). Bivariate analyses indicated that men were slightly older than women and had a lower social approval score. Although men had a higher baseline intake of energy, fat, and saturated fat from both 7DDR and 24HR in comparison with women, nutrient differences disappeared after energy adjustment (table 2).

Results of the general linear model analyses conducted on data from males, shown in table 3, indicate that social approval score and baseline BMI were significantly associated with increased fat, saturated fat, and energy scores, as estimated by the 7DDR relative to the 24HR. One unit increase in the social approval score was associated with an overestimate on the 7DDR relative to the 24HR of about 22 kcal/day in total energy intake and 1.2 g/day in total fat intake, and 1 kg/m^2 increase in BMI was related to an overestimate of 27 kcal/day of energy and 1.5 g/day of total fat intake as estimated by the 7DDR relative to the 24HR.

Results of general linear model analyses conducted on data from females are shown in table 4. It can be seen that a higher social desirability score was related to an underestimate in fat and energy intakes based on the 7DDR; and one unit increase in the social desirability score was associated with an underestimate of approximately 19 kcal/day in energy intake and 0.8 g/day in total fat.

Because a relatively large number of subjects (46 males and 78 females) did not have measured weight for use in computing BMI, we were concerned that omitting observations with missing values for BMI could distort results of the regression models. When BMI was omitted from these models, the effects of social approval on nutrient scores among males were attenuated. For example, we found that the regression coefficients describing the association between social approval score and 7DDR-derived nutrient scores were reduced as follows: for total fat, $\beta = 0.78$, standard error of beta (SE$_\beta$) = 0.40, and $p = 0.05$; for total saturated fatty acids, $\beta = 0.28$, SE$_\beta$ = 0.14, and $p = 0.04$; and for total energy, $\beta = 13.04$, SE$_\beta$ = 7.03, and $p = 0.06$. Interestingly, in females there was no change in effect when BMI was omitted from the models.

TABLE 1. Selected characteristics of responders and nonresponders at baseline, WATCH* Study, Worcester, Massachusetts, 1991–1995

Variables	Responders		Nonresponders		*p* value†
	No.	%	No.	%	
Categorical					
Gender					
Female	434	57.18	262	50.48	
Male	325	42.82	257	49.52	0.02
Marital status					
Married	533	70.22	328	63.20	
Others	226	29.78	191	36.80	0.01
Race					
White	680	89/59	393	75.72	
Others	79	10.41	126	24.28	<0.0001
Education					
No college	292	41.24	201	47.74	
Some college	413	58.58	220	52.26	0.04
Employment status					
Full-time	381	54.27	276	64.79	
Part-time	110	15.67	45	10.56	
No	211	30.06	105	24.65	0.002
Cigarette smoking					
Yes	120	18.21	88	22.22	
No	539	81.79	308	77.78	0.11
Plan to eat less red meat					
Yes	329	46.80	210	50.59	
No	374	53.20	215	49.41	0.40
Plan to eat less fat					
Yes	518	73.68	316	74.35	
No	185	26.32	109	25.65	0.80
Use exercise to lose weight					
Yes	164	31.42	84	27.63	
No	358	68.58	220	72.37	0.25

	Mean (SD*)	Mean (SD)	*p* value†
Continuous			
Age (years)	49.44 (10.63)	48.29 (10.45)	0.06
BMI*	28.80 (5.29)	29.51 (5.90)	0.05
Serum total cholesterol (mg/dl)	233.60 (46.85)	223.48 (56.59)	0.0008
Serum LDL* cholesterol (mg/dl)	152.82 (39.39)	148.89 (43.19)	0.10
No. of extra lipid measures	2.15 (1.26)	1.92 (1.22)	0.01
No. of visits to physician	3.32 (3.03)	3.45 (3.56)	0.51

* WATCH, Worcester Area Trial for Counseling in Hyperlipidemia; SD, standard deviation; BMI, body mass index (weight (kg)/height (m)^2); LDL , low density lipoprotein.

† *p* value based on chi-square test for categorical variables and two-sample *t* test of the difference between two group means.

DISCUSSION

Social desirability appeared to bias the dietary data of women, but not of men. Conversely, social approval bias was evident in men only. These biases were evident in comparing the raw nutrient scores and total energy intake derived from the 7DDR with the 24HR, but not in comparing the energy-adjusted nutrient scores. It is important to note that fat is the major determinant of variation in energy intake, in this study accounting for 70 and 83 percent of energy intake in the 24HR and 7DDR data, respectively. By controlling for energy, we may be factoring out the influence of fat as its major determinant. Therefore, in terms of assessing bias, it may be conceptually cleaner to focus on nutrient consumption without controlling for energy. This discussion focuses on those results.

Social desirability and social approval are response sets that have been shown to produce biases in self-report in a variety of testing situations (8, 10–13). The degree to which such biases are evident depends on beliefs and attitudes about the subject matter and ease with which respondents can identify responses as ei-

TABLE 2. Descriptive analyses of responders by gender at baseline, WATCH* Study, Worcester, Massachusetts, 1991–1995

	Male (n = 325)		Female (n = 434)		p value†
	Mean	SD*	Mean	SD	
Age (years)	50.37	10.61	48.75	10.60	0.04
Social desirability score	18.31	5.85	19.00	5.67	0.10
Social approval score	36.72	7.56	38.97	9.41	0.0003
BMI* (kg/m²)	28.65	4.50	28.91	5.84	0.54
Nutrient intake from 7-day diet recall					
Total energy (kcal/day)	2,085.63	927.83	1,786.32	752.11	0.0001
Total fat (g/day)	88.22	50.94	77.49	41.76	0.0036
Total SFA* (g/day)	29.39	17.66	25.25	13.50	0.0009
Total fat (% energy)	37.11	8.96	38.17	8.73	0.12
Total SFA (% energy)	12.38	3.78	12.55	3.44	0.52
Nutrient intake from 24-hour diet recall					
Total energy (kcal/day)	2,097.78	788.39	1,542.83	562.63	0.0001
Total fat (g/day)	73.16	40.85	56.28	30.05	0.0001
Total SFA (g/day)	26.12	16.16	19.35	11.31	0.0001
Total fat (% energy)	30.73	9.42	31.93	9.37	0.10
Total SFA (% energy)	10.93	4.24	11.04	4.40	0.73

* WATCH, Worcester Area Trial for Counseling in Hyperlipidemia; BMI, body mass index (weight (kg)/height (m)²); SFA, saturated fatty acids.
† Based on two-sample *t* test of the difference between two group means.

ther "good" or "bad" (10, 12). The latter depends on the structure and specific content of the assessment instrument. For reasons summarized earlier, we theorized that a structured questionnaire would be more subject to these biases than would an open-ended, short, time-focused interview such as the 24HR.

The results of this study were in broad agreement with the theoretical factors that led us to predict a social desirability bias in 7DDR responses relative to the 24HR in women. However, the bias was only about one third as large as that observed in 27 women participating in an earlier study who completed two 7DDRs and agreed to be called seven times on randomly selected days for telephone-administered 24HRs (23). In this study, we observed a significant bias in males between 7DDR- and 24HR-derived energy and fat intakes by both social approval and BMI. It may be that on measures of behavior in general, men with a need for social approval would be inclined to report a high level of the behaviors being asked about. That is the usual effect of an acquiescent response set. This pattern is consistent with the data for men (even the overweight men) in this sample. However, the expected acquiescent response in women appears to be overwhelmed by a tendency to report diet in a defensive way (as measured by social desirability) and thus underreport energy and fat intake. It is possible that the high-calorie/high-fat foods typically preferred by men (e.g., meats) are more likely to be seen as "good foods" by men (even in a study to lower cholesterol)

and thus to be overreported (28, 36). Food choices of men may be based more on traditional, outmoded notions of healthfulness and the primary comfort effects of high-fat foods (37). Either of these interpretations is possible, and other studies are needed to clarify these interesting gender differences.

Because of its sample size, results generated from the present study may be more stable than those obtained in the smaller study reported earlier (23); however, they also reflect differences in the WATCH population characteristics, representing both real and perceived needs to lower blood cholesterol levels. Consistent with the more intensive data collection procedures and the fact that subjects had to volunteer without referral for a medical condition in the previous study (23), there was a higher response rate to the questionnaires (90 percent) and a shorter response interval (9.0 days on average) than was observed in this study (i.e., 59.4 percent overall and 81.3 percent response for those with complete dietary data and response interval averaging 22.6 days), reflecting marked differences between subjects in the two studies.

In this study, comparison data were derived from a single 24HR, as opposed to seven 24HRs in the prior study. It could be that the first 24HR is more subject to the influence of social desirability, and these defensive motivations are attenuated in subsequent administrations of the 24HR. Also, it is known that due to the large intraperson variability in dietary intake, especially day to day, a single 24HR will not adequately

TABLE 3. Results of the general linear models to assess the effect of social desirability and social approval bias in nutrient estimate for males (n = 224), WATCH* Study, Worcester, Massachusetts, 1991–1995†

Dependent variables	Independent variables			
	24-hour-derived nutrient score	Social desirability score	Social approval score	BMI*
Total fat (g/day)				
β‡	0.29	0.70	1.21	1.48
p§	0.0003	0.24	0.01	0.04
Total SFA* (g/day)				
β‡	0.26	0.25	0.45	0.50
p§	0.0001	0.23	0.01	0.05
Total energy (kcal/day)				
β‡	0.36	9.18	21.50	27.24
p§	0.0001	0.38	0.01	0.03
Total fat (% energy)				
β‡	0.25	0.09	0.10	0.09
p§	0.0002	0.43	0.22	0.49
Total SFA (% energy)				
β‡	0.30	0.02	0.05	0.06
p§	0.0002	0.60	0.09	0.23

* WATCH, Worcester Area Trial for Counseling in Hyperlipidemia; BMI, body mass index (weight (kg)/height (m)²); SFA, saturated fatty acids.

† All models were fitted with the 7-day diet recall-derived nutrient as dependent variable, corresponding 24-hour-derived nutrient, social desirability score, social approval score, body mass index (kg/m²), interval (in days of mailing and receipt of the questionnaires), and incentive as independent variables.

‡ Regression coefficient obtained in the model, computed using PROC GLM in SAS (34) and based on the type III, orthogonal sums of squares (i.e., with the variable listed as though it were entered last). It is the estimated effect for the variable listed. For example, for each point in the social approval score, we estimate a 1.21-g/day increase in fat consumption derived from 7-day diet recall relative to 24-hour-derived fat intake.

§ Value of the test of H_0: β = 0.

represent an individual's dietary intake (38). To some extent, this countervails the benefit of having a much larger sample size. It is widely known that such a large contribution of intraperson variability to overall variance attenuates the agreement between nutrient scores derived from two instruments (39).

Besides time frame, the 24HR and the 7DDR differ in respect to mode of administration, the former being interviewer-administered and the latter, self-administered. There is some evidence that expert interviewers can elicit a defensive response in interviewees, especially for behaviors that have strong social connotations (40) or for outcomes that are particularly onerous for which the interviewer is aware of study hypotheses (41). That results derived from interviewer-based methods differ by age and education (9) indicates a possible association between perceptions of the connection between diet and health status and the quality of data collected. Besides the methodological explanations for the lower social desirability bias among women in this study, these results are consistent with another study in which individuals who were in interventions to reduce cardiac risk factors were found to be less judgmental or defensive regarding

their dietary lapses than those in the general population who were making changes for weight loss (26).

In our study, there was a sizable number of nonresponders. We found that these individuals differed from responders on important factors known to affect food intake, including gender, marital status, race, and BMI. Previously we had found that delay in response (perhaps a rough proxy for nonresponse) was associated with 24HR- and 7DDR-derived nutrient score differences (23). The fact that nonresponders differed with respect to factors found to bias such measures (i.e., gender and BMI) calls into question biases that may exist in epidemiologic studies in which data collection routines vary in terms of participant demand and factors related to self-selection. A recent study of cognitive distortions regarding nutritional information found extensive evidence for judgmental attitudes and good/bad dichotomous thinking about food according to age, education, health consciousness, and socioeconomic status (42). Unfortunately, gender differences were not analyzed in that study. The fact that subjects appear to think of fat in foods almost as though it were an infectious agent raises interesting questions about

TABLE 4. Results of the general linear models to assess the effect of social desirability and social approval bias in nutrient estimate for females (n = 291), WATCH* Study, Worcester, Massachusetts, 1991–1995†

Dependent variables	Independent variables			
	24-hour-derived nutrient value	Social desirability score	Social approval score	BMI*
Total fat (g/day)				
β‡	0.28	−0.78	−0.01	0.52
p§	0.0002	0.07	0.98	0.18
Total SFA* (g/day)				
β‡	0.26	−0.26	0.02	0.15
p§	0.0001	0.05	0.78	0.24
Total energy (kcal/day)				
β‡	0.27	−19.16	0.02	8.55
p§	0.0002	0.02	0.99	0.22
Total fat (% energy)				
β‡	0.27	0.00	0.00	0.13
p§	0.0001	0.98	0.95	0.14
Total SFA (% energy)				
β‡	0.26	−0.01	0.01	0.05
p§	0.0001	0.69	0.75	0.14

* WATCH, Worcester Area Trial for Counseling in Hyperlipidemia; BMI, body mass index (weight (kg)/height (m)^2); SFA, saturated fatty acids.

† All models were fitted with the 7-day diet recall-derived nutrient as dependent variable, corresponding 24-hour-derived nutrient, social desirability score, social approval score, body mass index (kg/m^2), interval (in days of mailing and receipt of the questionnaires), and incentive as independent variables.

‡ Regression coefficient obtained in the model, computed using PROC GLM in SAS (34) and based on the type III, orthogonal sums of squares (i.e., with the variable listed as though it were entered last). It is the estimated effect for the variable listed. For example, for each point in the social desirability score, we estimate a 19.16-kcal/day decrease in energy consumption derived from 7-day diet recall relative to 24-hour-derived fat intake.

§ Value of the test of H_0: β = 0.

perceptions of subjects around food intake relative to epidemiologists' need to estimate nutrient effects.

As with most studies of diet, this study is limited by the lack of a true criterion measure. The 24HR is used as the comparison measure in this study, but it also is subject to a variety of potential biases. Doubly labeled water, a biologic measure for total energy intake that is not subject to social desirability and approval biases, might be used in future studies (43) as might urinary nitrogen for protein intake or 3-methylhistidine for meat intake (44). These comparison methods, however, tend to be very costly and cumbersome in terms of biological sample collection. As we consider ways to improve dietary assessment, we may need to broaden our conceptualization of validity to include constructs such as weight or serum cholesterol measurements. These parameters relate changes in overall energy balance or specific nutrient exposures according to well-established physical or biological laws (or both) (45–48). If our dietary assessment instruments are measuring what we intend them to, then their results should reconcile against such parameters, at least on a group level (e.g., for cholesterol) (49).

The 7DDR was developed to assess the relation between changes in diet and blood lipids, which tend to occur over a short period. It is similar to the FFQ in its use of a finite list of foods, but it is less reliant on generic memory of food intake. From a theoretical perspective, the FFQ, with its focus on the longer term, is perhaps more prone to bias than is the 7DDR. Future work in this field should assess the FFQ for such biases. Comparison assessments with different modes of administration can provide additional clues about the nature of these biases.

CONCLUSION

Future research should focus on examining the potential mediators of the observed gender-specific biases. Dieting status alone might account for much of the effect observed in women. Other mediating variables could include nutritional knowledge, obesity, binge and stress-induced eating, and degree of endorsement of beliefs that certain foods are "good" or "bad" (beliefs often rigidly held by those high in dietary restraint). Factors affecting bias in dietary self-report in men, though less clearly conceptualized when we began this study, are equally important to identify and measure.

Biases due to social desirability or social approval can confound the relation between diet and disease. In case-control studies, knowledge of disease status could condition cases' and controls' perceptions of the diet-disease relation, thus setting the stage for bias. Even in follow-up studies, confounding can occur if either social desirability or social approval biases estimates of nutrient exposure. Psychological attributes associated with social desirability or social approval (i.e., an acquiescent personality type) are associated with physiologic processes (50) that are, in turn, related to disease outcomes (51).

Given that studies of the relations between nutrition and disease are becoming more prominent and that such studies often provide equivocal results, it is important to determine whether bias is contributing to some distortion in those results. If social desirability and social approval biases prove to be large and predictable, they will need to be controlled in the analyses of dietary data to assess accurately the relation between self-reported nutrient intakes and disease. Additional studies are needed to establish models to adjust for these biases in different populations.

ACKNOWLEDGMENTS

REFERENCES

1. Dwyer JT, Krall EA, Coleman KA. The problem of memory in nutritional epidemiology research. J Am Diet Assoc 1987; 87:1509–12.
2. Smith AF, Jobe JB, Mingay DJ. Question-induced cognitive biases in reports of dietary intake by college men and women. Health Psychol 1991;10:244–51.
3. Smith AF. Cognitive psychological issues of relevance to the validity of dietary reports. Eur J Clin Nutr 1993;47(Suppl 2):S6–18.
4. Subar AF, Frey CM, Harlan LC, et al. Differences in reported food frequency by season of questionnaire administration: The 1987 National Health Interview Survey. Epidemiology 1994; 5:226–33.
5. Jain M, Howe GR, Harrison L, et al. A study of repeatability of dietary data over a seven-year period. Am J Epidemiol 1989;129:422–9.
6. Jensen OM, Wahrendorf J, Rosenquist A, et al. The reliability of questionnaire-derived historical dietary information and temporal stability of food habits in individuals. Am J Epidemiol 1984;120:281–90.
7. Byers T, Rosenthal RI, Marshall JR, et al. Dietary history from the distant past: a methodological study. Nutr Cancer 1987;5: 69–77.
8. Anastasi A. Psychological testing. 6th ed. New York, NY: MacMillan Publishing Co., 1988.
9. Dolan P, Kind P. Inconsistency and health state valuations. Soc Sci Med 1996;42:609–15.
10. Edwards AL. Social desirability or acquiescence in the MMPI? A case study with the SD scale. J Abnormal Soc Psychol 1961;63:351–9.
11. Edwards A. The social desirability variable in personality assessment and research. New York, NY: Dryden, 1957.
12. Marlowe D, Crowne DP. Social desirability and responses to perceived situational demands. J Consult Clin Psychol 1961; 25:109–15.
13. Larsen KS, Martin HJ, Ettinger RH, et al. Approval seeking, social cost, and aggression: a scale and some dynamics. J Psychol 1976;94:3–11.
14. Krauchi K, Wirz-Justice A, Graw P. The relationship of affective state to dietary preference: winter depression and light therapy as a model. J Affect Disorders 1990;20:43–53.
15. Roering KJ, Boush DM, Shipp SH. Factors that shape eating patterns: a consumer perspective. What is America eating? Washington, DC: National Academy Press, 1986:72–84.
16. Suitor CJW, Gardner J, Willett WC. A comparison of food frequency and diet recall methods in studies of nutrient intake of low-income pregnant women. J Am Diet Assoc 1989;89: 1786–94.
17. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122:51–65.
18. Smith AF. Cognitive processes in long-term dietary recall. Washington, DC: US GPO, 1991. (DHHS publication no. 4, series 6 (PHS) 91 1079).
19. Thompson FE, Metzner HL, Lamphiear DE, et al. Characteristics of individuals and long term reproducibility of dietary reports: the Tecumseh Diet Methodology Study. J Clin Epidemiol 1990;43:1169–78.
20. Friedenreich CM. Measurement of past diet: review of previous and proposed methods. Epidemiol Rev 1992;14:177–96.
21. Wu ML, Whittemore AS, Jung DL. Errors in reported dietary intakes. I. Short-term recall. Am J Epidemiol 1986;124: 826–35.
22. Jobe JB, Mingay DJ. Cognitive research improves questionnaires. Am J Public Health 1989;79:1053–5.
23. Hebert JR, Clemow L, Pbert L, et al. Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. Int J Epidemiol 1995;24:389–98.
24. Wardle J, Beales S. Restraint, body image and food attitudes in children from 12 to 19 years. Appetite 1986;7:209–17.
25. Horm J, Anderson K. Who in America is trying to lose weight? National Health Interview Survey. Proceedings of methods for voluntary weight loss and control: an NIH technology assessment conference. Hyattsville, MD: National Center for Health Statistics, Centers for Disease Control and Prevention, 1992:50–7.
26. Jeffrey RW, French SA, Schmid TL. Attributions for dietary failures: problems reported by participants in the hypertension prevention trial. Health Psychol 1990;9:315–29.
27. Klem ML, Klesges RC, Bene CR, et al. A psychometric study of restraint: the impact of race, gender, weight, and marital status. Addict Behav 1990;15:147–51.
28. Seim HC, Fiola JA. A comparison of attitudes and behaviors of men and women toward food and dieting. Fam Pract Res J 1990;10:57–63.
29. Johnson WG, Corrigan SA, Schlundt DG, et al. Dietary restraint and eating behavior in the natural environment. Addict Behav 1990;15:285–90.
30. Ockene IS, Hebert JR, Ockene JK, et al. Effect of training and a structured office practice setting on physician-delivered nutrition counseling : the Worcester Area Trial for Counseling in Hyperlipidemia (WATCH). Am J Prev Med 1996;12:252–8.
31. Martin HJ. A revised measure of approval motivation and its relationship to social desirability. J Pers Assess 1984;48: 508–16.
32. NDS, version 2.3. Food database, version 5A. Nutrient database, version 20. Minneapolis, MN: University of Minnesota Nutrition Coordinating Center, 1991.
33. Hebert JR, Ockene IS, Hurley TG, et al. Development and testing of a seven-day dietary recall. J Clin Epidemiol 1997; 50:925–37.

34. SAS. SAS user's guide: statistics, version 6.11. Cary, NC: SAS Institute, Inc., 1996.
35. SAS. SAS user's guide. Cary, NC: SAS Institute, Inc., 1996.
36. Herman CP. Restrained eating. In: Stunkard A, ed. The psychiatric clinics of North America. Philadelphia, PA: Saunders, 1978
37. Drewnowski A. Dietary fats: perceptions and preferences. J Am Coll Nutr 1990;9:431–5.
38. Willett WC. Nutritional epidemiology. Oxford, England: Oxford University Press, 1990.
39. Rosner B, Willett WC. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. Am J Epidemiol 1988;127:377–86.
40. Aquilino W. Telephone versus face-to-face interviewing for household drug use surveys. Int J Addict 1992;27:71–91.
41. Hertz-Picciotto I, Swan SH, Neutra RR. Reporting bias and mode of interview in a study of adverse pregnancy outcomes and water consumption. Epidemiology 1992;3:104–12.
42. Rozin P, Ashmore M, Markwith M. Lay American conceptions of nutrition: dose insensitivity, categorical thinking, contagion, and the monotonic mind. Health Psychol 1996;15:438–47.
43. Bingham SA. The use of 24-h urine samples and energy expenditure to validate dietary assessments. Am J Clin Nutr 1994;59(suppl):227S-31S.

44. Hebert JR, Miller DR. Methodologic considerations for investigating the diet-cancer link. Am J Clin Nutr 1988;47:1068–77.
45. Keys A, Anderson JT, Grande F. Serum cholesterol response to changes in the diet. III. Differences among individuals. Metabolism 1965;14:766–75.
46. Hegsted DM, McGandy RB, Myers ML, et al. Quantitative effects of dietary fat on serum cholesterol in man. Am J Clin Nutr 1965;17:281–95.
47. Flatt JP. Dietary fat, carbohydrate balance, and weight maintenance: effects of exercise. Am J Clin Nutr 1987;45:296–306.
48. Romieu I, Willett WC, Stampfer MJ, et al. Energy intake and other determinants of relative weight. Am J Clin Nutr 1988;47:406–12.
49. Jacobs DJ Jr, Anderson JT, Blackburn H. Diet and serum cholesterol: Do zero correlations negate the relationship? Am J Epidemiol 1979;110:77–87.
50. Fox B. Psychogenic factors in cancer, especially its incidence. In: Maes S, Spielberger C, Defares P, et al, eds. Topics in health psychology. New York, NY: John Wiley & Sons, 1988:37–55.
51. Kune GA, Kune S, Watson LF, et al. Personality as a risk factor in large bowel cancer: data from the Melbourne Colorectal Cancer Study. Psychol Med 1991;21:29–41.