International Journal of Epidemiology
© International Epidemiological Association 1995
Vol. 24, No. 2
Printed in Great Britain

Social Desirability Bias in Dietary Self-Report May Compromise the Validity of Dietary Intake Measures

JAMES R HEBERT,* LYNN CLEMOW,* LORI PBERT,* IRA S OCKENE† AND JUDITH K OCKENE*

Hebert J R (Division of Preventive and Behavioral Medicine, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, USA), Clemow L, Pbert L, Ockene I S and Ockene J K. Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. *International Journal of Epidemiology* 1995; **24:** 389–398.

Background. Self-report of dietary intake could be biased by social desirability or social approval thus affecting risk estimates in epidemiological studies. These constructs produce response set biases, which are evident when testing in domains characterized by easily recognizable correct or desirable responses. Given the social and psychological value ascribed to diet, assessment methodologies used most commonly in epidemiological studies are particularly vulnerable to these biases.

Methods. Social desirability and social approval biases were tested by comparing nutrient scores derived from multiple 24-hour diet recalls (24HR) on seven randomly assigned days with those from two 7-day diet recalls (7DDR) (similar in some respects to commonly used food frequency questionnaires), one administered at the beginning of the test period (pre) and one at the end (post). Statistical analysis included correlation and multiple linear regression.

Results. Cross-sectionally, no relationships between social approval score and the nutritional variables existed. Social desirability score was negatively correlated with most nutritional variables. In linear regression analysis, social desirability score produced a large downward bias in nutrient estimation in the 7DDR relative to the 24HR. For total energy, this bias equalled about 50 kcal/point on the social desirability scale or about 450 kcal over its interquartile range. The bias was approximately twice as large for women as for men and only about half as large in the post measures. Individuals having the highest 24HR-derived fat and total energy intake scores had the largest downward bias due to social desirability.

Conclusions. We observed a large downward bias in reporting food intake related to social desirability score. These results are consistent with the theoretical constructs on which the hypothesis is based. The effect of social desirability bias is discussed in terms of its influence on epidemiological estimates of effect. Suggestions are made for future work aimed at improving dietary assessment methodologies and adjusting risk estimates for this bias.

Methods of collecting dietary data used in most epidemiological studies may be prone to biases stemming from the respondent's wishes to convey a desirable image or to seek approval for certain behaviours. Social desirability is the tendency of an individual to convey an image in keeping with social norms and to avoid criticism in a 'testing' situation.[1–7] Social approval is the tendency for an individual to seek a positive response in the testing situation and is therefore less focused on defensiveness.[8–11] In general, scores from the two scales are only weakly positively correlated.[9] The response biases that result from social desirability or social approval can significantly obscure or distort the measurement of the variable of interest. Therefore, estimating these biases on a particular questionnaire is a necessary precondition to the adoption and use of a measure in psychological practice or research.[3–9,12]

The connection between diet and both general health status and specific diseases has been widely popularized, especially in the West.[13,14] For example, an increasing number of food products in the US are sold with a health claim (or disclaimer) either explicitly stated or implied.[15,16] Consequently, it is now widely held that specific types of foods or patterns of eating are healthy, or desirable, and others are unhealthy, or undesirable.[17]

It is also apparent that specific dietary recommendations aimed at population subgroups appear to influence dietary self-report. For example, pregnant women have been shown to upward bias their estimates of total energy intake in a manner consistent with antenatal dietary advice.[18] In other female populations with

* Division of Preventive and Behavioral Medicine, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, USA.

† Division of Cardiovascular Medicine, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, USA.

generally opposite demand characteristics (i.e. where lower intakes are the ideal) a downward bias is observed.[19,20] Even when the food frequency questionnaire (FFQ) method produces larger overall estimates of energy intake, the overestimates for nutrients associated with undesirable foods (e.g. fat) are generally proportionally smaller.[21]

Given the wide public knowledge concerning nutritional guidelines for health, self-report of eating behaviour is extremely vulnerable to social desirability bias. Despite some evidence that social desirability scores are related to the reported consumption of fresh fruits and vegetables and snack foods,[22] dietary assessment questionnaires have not been checked rigorously for response bias.

Instruments such as the FFQ, which aim to establish long-term habitual intake, are the methods of choice for dietary assessment in larger-scale studies. The potential bias in these methods is greater than for short-period assessments such as food diaries (FD) or 24-hour diet recalls (24HR) in which the participant is asked about intake of specific foods either while they are being eaten or during a well-defined period in the very recent past. Although it is recognized that there are gender, age, and other differences in dietary self-report,[18,21,23–26] it is generally assumed that the typically >75% unexplained variability in FFQ-derived nutrient scores in relation to the typical 24HR or FD falls in the category of intraperson sources which are regarded as random. However, problems even more serious than attenuation of the risk estimate may exist if the error in measurement is not random, i.e. if it is systematically biased due to response set biases such as those due to social desirability and social approval.

Because of obvious gender differences in food behaviour[27] and emotional issues around food,[28] one might expect gender differences in social desirability or approval biases. Issues such as self- and public attention,[29,30] premenstrual syndrome,[31] depression,[32] and obesity[33] tend to be positively associated with the intake of certain foods (often high fat/high carbohydrate foods), chosen possibly more for their 'comfort' qualities[17] than their nutritional content. Some of these factors, especially obesity[23,34,35] and restrained eating[36] are known to differ by gender. Estimates vary, but recent figures[37] suggest that between 33 and 40% of US women are currently dieting, nearly twice the rate of men, and an additional 28% are trying to maintain a weight loss. Women seem to be more vulnerable to stress-induced eating than men.[38] Dieting, guilt about eating, restrained eating, and binge eating are all twice as common in women as men[30,39,40] and may further distort self-report of food intake in women. Factors that have strong emotional import may have particular relevance to dietary self-report because current knowledge of neurophysiology informs us that the same portion of the brain that performs olfactory functions also has an important role in emotion and instinctual behaviour.[41]

METHODS

In October–November of 1991, we conducted a validation study of a 7-day diet recall (7DDR) developed to assess fat intake in the Worcester Area Trial for Counseling in Hyperlipidemia (WATCH).[42] The 7DDR was filled out twice, once at the beginning of the 3-week test period (pre measures) and once at the end (post measures). Nutrient scores (i.e. estimated daily intake) derived from the 7DDR were compared to those derived from seven 24HR randomly administered over the 3 weeks such that each day of the week was represented once. We chose the 24HR because of the low variance estimates obtained in comparing this method to records and to frequency methods.[43–45]

The 7DDR looks very much like an FFQ. However, it is dissimilar in that it attempts to elicit a response regarding specific food encounters and not long-term, habitual intake. To the extent that recall of a specific episode is not possible, subjects are instructed to provide their usual consumption, as they would on an FFQ. Indeed, correlations between fat-related nutrients from the 7DDR and the 24HR (usually >0.70) were higher than those which are typically obtained from FFQ. The slope of the regression lines obtained by regressing 7DDR-derived nutrient values on 24HR-derived values did not differ significantly from 1.0.[42]

In October 1993, exactly 2 years after the study had ended, we administered both a Marlow-Crowne Social Desirability Scale (MCSD)[7] and a Martin-Larsen Social Approval Scale (MLSA)[9] to each of the 41 validation study participants who had completed at least one of the two 7DDR. The MCSD consists of 33 true-false questions (e.g. 'I never hesitate to go out of the way to help someone in trouble'). Eighteen questions (including this example) are scored one point on a 'true' response and 15 are scored one point on a 'false'. The MLSA consists of 20 questions requiring 5-point Likert scale responses ranging from 'disagree strongly' to 'agree strongly' (e.g. 'I seldom feel the need to make excuses or apologize for my behaviour'). Items are rated on a 5-point scale with 'agree strongly' rating 5 points except for five questions which are reverse-scored (including the example).

Analysis of these data included simple univariate statistics, correlation, and linear regression. For this we

TABLE 1 *Descriptive statistics including daily intake of selected nutritional variables by sex based on multiple 24-hour diet recalls, WATCH Nutrition Validation and Social Desirability Studies, Worcester, MA, USA, 1991–1993*

Variable	Mean/SD[a]			
	Female (n = 27)	Male (n = 14)	T/χ^2 [b]	*P*-value
% married	74.1%	64.3%	0.43[c]	0.72[c]
% white	81.5%	85.7%	0.12[c]	1.00[c]
Age (years)	51.3/16.1	47.2/13.9	–0.81	0.42
Education (years)	13.4/1.9	13.5/2.1	0.17	0.86
BMI[d]	26 6/5.6	20.6/3.4	0.53	0.60
Social desirability score	20.5/6.4	18.5/5.9	0.89	0.38
Social approval score	48.4/9.6	44.9/9.1	1.01	0.32
Total energy (kcal)	1490/461	1970/599	–2.85	0.007
Total fat (g)	51.9/17.3	65.1/23.8	–2.05	0.05
Saturated fatty acids (g)	18.2/7.0	21.6/8.7	–1.35	0 18
Monounsaturated fatty acids (g)	19.0/6.6	25.0/10.3	–2.29	0.03
Polyunsaturated fatty acids (g)	10.5/3.5	13.2/5.5	–1.89	0.07
% calories from fat	31.3/5.5	29.4/4.6	1.10	0.28
Cholesterol (mg)	185.1/64.8	261.5/194.5	–1.43	0.17
Alcohol (g)	3 9/8.5	15.5/17.9	–2.30	0.04
Dietary fibre (g)	14.1/5.7	16.5/10.4	–0.80	0.43

[a] Standard Deviation.

[b] Difference by sex is based on T-test for each demographic or nutrient/nutritional parameter shown unless it is a test of a categorical variable, in which case it is the χ^2 statistic or Fisher's Exact test and is noted by [c].

[c] Based on the χ^2 test of independence. The *P*-value for % white is based on Fisher's Exact test owing to the presence of fewer than 5 observations in some cells.

[d] Body mass index = weight (kg)/height (m)^2.

used PROCS UNIVARIATE, CORR and GLM in SAS.[46] All regression analyses accounted for age, gender, and body mass index, variables known or thought to be related to differences in dietary intake.[23,37,39,40] Besides fitting either social desirability or social approval scores in the model, we also fitted interval (in days) between mailing of the forms and their return because we thought that this would be an estimate of true (as opposed to perceived) compliance, which lies close to the theoretical domains of social approval and social desirability. In addition to the overall models, we stratified analyses by gender.

We also sought to determine if social desirability bias differed according to level of nutrient intake, here estimated by the 24HR-derived scores. For this we did a separate set of analyses stratifying by quartiles of the two nutrients for which we observed the largest overall effect of social desirability. We conducted all analyses using both pre and post measures; the former because these correspond to the 7DDR measures most likely encountered in an epidemiological study (i.e. unaffected by training from responding to multiple diet assessments) and the latter because responses to the second 7DDR are concordant with the period during which the multiple 24HR were collected.

RESULTS

Descriptive analyses are shown in Table 1. In general, there was a wide range of variability in social desirability scores, with an interquartile range of nine points for both men and women but a wider overall range in women (7–33) than in men (11–28). The overall distribution of social approval also was smaller in men (32–57) than in women (28–67) but the interquartile range was larger in men (20 versus 13 points). In general, we observed that calorie-contributing components were consistently higher for men than for women, but there were no significant differences on other factors.

Bivariate analyses indicated that social approval was not related to any nutritional variable (i.e. $r < |0.22|$, $P > 0.19$ in every instance with an average of $r = -0.08$). In contrast, social desirability was highly correlated with a large number of nutritional variables in both the pre and post measures with r averaging

−0.31 and −0.32, respectively. That the patterns of correlation persist across the nutrients should not be surprising because nutritional variables tend to be very highly intercorrelated (r averaging approximately 0.85). Correlations tended to be more strongly negative for 7DDR- than for 24HR-derived scores; for example, for pre measures correlations for energy were $r = -0.47$ ($P = 0.004$) and $r = -0.31$ ($P = 0.06$) for 7DDR and 24HR, respectively. Post measures produced nearly identical correlations: i.e. $r = -0.50$ ($P = 0.003$) and $r = -0.31$ ($P = 0.06$), respectively. In all sex-stratified analyses females had consistently more strongly negative correlations; for energy in the pre measures for the 7DDR and 24HR derived scores $r = -0.57$ ($P = 0.003$) and $r = -0.39$ ($P = 0.05$), respectively. For males, correlations were much closer to 0. For example, for men the corresponding correlations for energy were $r = -0.22$ ($P = 0.52$) and $r = -0.03$ ($P = 0.92$), respectively. We did observe a number of significant correlations between social desirability score and the difference of the 7DDR-derived and 24HR-derived nutrient scores. These included: total energy intake in both the pre ($r = -0.42$, $P = 0.001$) and post ($r = -0.40$, $P = 0.02$); non-fat calories in both the pre ($r = -0.39$, $P = 0.02$) and post ($r = -0.38$, $P = 0.03$); and for cholesterol in the pre measures only ($r = -0.34$, $P = 0.04$). Social desirability and social approval scores were not correlated with each other ($r = 0.09$, $P = 0.60$), consistent with their strikingly different patterns of correlation with the nutritional variables and reflecting the fact that they encompass different psychological constructs.[9]

Results from the general linear models are shown in Table 2 for pre measures, and Table 3 for post measures. In general, neither age nor sex were significant predictors of 7DDR-derived score after accounting for 24HR-derived score and interval. Therefore, models shown exclude age and sex but include social desirability score, interval (between our mailing the form and the respondent returning it), and body mass index. All models were re-run stratified by gender. General linear models using difference scores as the dependent variable as opposed to the models shown yielded virtually identical results. The models shown were chosen because they allow for direct estimation of the regression of 7DDR-derived score on 24HR-derived score. As for the results of the correlation analyses, social approval score was not a significant predictor in any model.

For nearly all of the variables shown there was a significant or marginally significant downward bias of 7DDR-derived nutrient score with increasing social desirability score, especially for the pre measures. Without exception, the magnitude of the reduction was smaller in the post measures, with concomitant reduction in the level of significance. Body mass index generally was associated with a positive regression coefficient in these data. Interval between mailing and return of the social desirability and approval questionnaires exerted an effect generally in the same direction as that of social desirability and in some instances approaching it in magnitude.

As shown in Table 4, nearly without exception the downward bias due to social desirability in the 7DDR relative to the 24HR was greater for women than for men. The ratios of the female to male regression coefficients were generally between 2.0 and 3.0. Even though the relative difference between the genders was sometimes smaller in the pre measures, the absolute difference or bias due to gender was generally smaller in the post measures. By using regression models stratified by quartiles of the 24HR-derived nutrient score we obtained the results shown in Table 5. These show a pattern approximating a j-shaped curve in the pre measures, with first and third quartile values of approximately equal magnitude, second quartile values the lowest, and fourth quartile values about threefold higher than first and third quartile values. This pattern was only vaguely evident in the post measures.

DISCUSSION

This section first focuses on our results and potential relevance of the bias we detected in this study and then expands to the general matter of bias in dietary assessment methods including the FFQ.

Interpretation of the Results of this Study

Our results showed a relatively large bias due to social desirability that was, in general, consistent with our original hypothesis. The bias was generally larger in women than men and appeared to have decreased over the course of the study. However, the gender differences tended to persist. There was a suggestion that the bias increased with level of fat and total energy intake in the pre measures in a manner that was not evident in the post measures. This result is consistent with a decrease in reactivity with increased exposure to dietary assessment.

Although the bias that we observed would, theoretically, be strongest using a technique such as the FFQ, we had no feasible way of assessing this except by using the 7DDR because it was that instrument we used in the WATCH. The 7DDR is only moderately similar to the FFQ, in its grid-like ranking structure and its partial reliance on habitual memory. This issue remains to be

TABLE 2 *Results of the general linear models to assess social desirability bias in nutrient estimation, pre measures, WATCH Nutrition Validation and Social Desirability Studies, Worcester, MA, USA, 1991–1993*

		Independent Variables[a]			
		24-HR Score[b]	SD Score[c]	BMI[d]	Interval
Dependent[e] variable					
Total fat (g)	b[a] =	1.41	-2.34	1.39	-1.35
	SE_b[a] =	0.39	1.17	1.27	0.75
	P-value[a] =	n.s.	(0.06)	n.s.	(0.08)
Saturated fatty acids (g)		1.35	-0.68	0.63	-0.30
		0.29	0.32	0.35	0.20
		n.s.	(0.04)	(0.09)	n.s.
Monounsaturated fatty acids (g)		1.29	-0.92	0.54	-0.52
		0 36	0.44	0 48	0.28
		n.s.	(0.05)	n.s.	(0.08)
Polyunsaturated fatty acids (g)		0.72	-0.88	-0.01	-0.24
		0.56	0.38	0.42	0 26
		n s.	(0.03)	n s.	n.s.
Alcohol (g)		0.67	-0.17	-0.11	0.06
		0.08	0.15	0.18	0.10
		(< 0 0001)	n.s.	n.s.	n.s.
Total energy (kcal)		0.96	-50.23	33.68	-30.62
		0.27	22.58	25.17	14.57
		n.s.	(0.03)	n.s.	(0.05)
Non-fat energy (kcal)		0.78	-26.15	21.25	-19.88
		0 21	12.70	14.34	8.23
		n.s.	(0.05)	n.s.	(0.02)
Fat (% energy)		0.61	-0.29	0.21	0.07
		0.17	0.16	0.19	0.10
		(0.06)	(0.08)	n.s.	n.s.
Cholesterol (mg)		0.63	-7.77	6.63	-1.29
		0.20	2.34	2.56	1.48
		n.s.	(0.003)	(0.02)	n.s.
Dietary fibre (g)		0.25	-0.31	0.09	-0.16
		0.12	0.16	0.17	0.10
		(< 0.0001)	(0.05)	n.s.	n.s.

[a] All models were fitted with the 24-hour diet recall-derived nutrient variable corresponding to the 7-day diet recall (7DDR)-derived dependent variable, as well as social desirability score, body mass index, and interval (in days) between posting and return of the social desirability questionnaire as the independent variables. Each independent variable lists in column format the regression coefficient, standard error of the regression coefficient, and the (*P*-value) given as n.s. if $P > 0.10$ of the F-test of Ho:$\beta = 0$, except for 24-HR score, as noted below. All models are based on type III (orthogonal) sums of squares.

[b] This is the nutrient score derived from seven random 24-hour diet recalls corresponding to the dependent variable listed. Unlike for the other independent variables listed, here we test Ho:$\beta = 1.0$, because the ideal is complete agreement between the 24HR-derived and 7DDR-derived nutrient scores.

[c] SD score is the social desirability score as obtained from the Marlowe-Crowne Social Desirability Scale.[1]

[d] Body mass index = weight (kg)/height $(m)^2$.

[e] The nutrient values are in daily amounts specified or they are expressed as percentages.

TABLE 3 *Results of the general linear models to assess social desirability bias in nutrient estimation, post measures, WATCH Nutrition Validation and Social Desirability Studies, Worcester, MA, USA, 1991–1993*

		Independent Variables[a]			
		24-HR Score[b]	SD Score[c]	BMI[d]	Interval
Dependent[e] variable					
Total fat (g)	b =[a]	0.9	−1.18	1.86	−0.54
	SEb =[a]	0.20	0.59	0.64	0.39
	P- value =[a]	n.s.	(0.06)	(0.007)	n.s.
Saturated fatty acids (g)		0.87	−0 30	0.59	−0.18
		0.17	0.20	0.22	0.16
		n.s.	n.s.	(0.01)	n.s.
Monounsaturated fatty acids (g)		0.93	−0.54	0.68	−0.19
		0.20	0 24	0.27	0.16
		n s.	(0 04)	(0.02)	n.s.
Polyunsaturated fatty acids (g)		0.75	−0.37	0.45	−0.06
		0.23	0.15	0.17	0.10
		n.s.	(0 02)	(0.01)	n.s.
Alcohol (g)		0.96	−0.12	−0 04	−0.05
		0.12	0.22	0.27	0.15
		n.s	n.s.	n.s.	n.s.
Total energy (kcal)		0.97	−25.50	31.48	−20.49
		0.17	14.49	16.16	9.37
		n.s	(0.09)	(0.06)	(0.04)
Non-fat energy (kcal)		0.89	−16.28	13.42	−13.93
		0.15	9.68	10.93	6.28
		n.s.	(0.10)	n.s.	(0.03)
Fat (% energy)		1.11	0.01	0.34	0.09
		0.18	0.17	0.20	0.11
		n.s	n.s.	(0.10)	n.s.
Cholesterol (mg)		0.52	−2.30	3.59	−1.81
		0.19	2.15	2.35	1.36
		(0.01)	n.s.	n.s.	n.s.
Dietary fibre (g)		0.52	−0.31	0.07	−0.11
		0.10	0.14	0.16	0.09
		(< 0.0001)	(0.04)	n.s.	n s.

[a] All models were fitted with the 24-hour diet recall-derived nutrient variable corresponding to the 7 day diet recall (7DDR)-derived dependent variable, as well as social desirability score, body mass index, and interval (in days) between posting and return of the social desirability questionnaire as the independent variables. Each independent variable lists in column format the regression coefficient, standard error of the regression coefficient, and the (*P*-value) given as *n.s.* if $P > 0.10$ of the F-test of Ho:$\beta = 0$, except for 24-HR score, as noted below. All models are based on type III (orthogonal) sums of squares.

[b] This is the nutrient score derived from seven random 24-hour diet recalls corresponding to the dependent variable listed. Unlike for the other independent variables listed, here we test Ho:$\beta = 1.0$, because the ideal is complete agreement between the 24HR-derived and 7DDR-derived nutrient scores.

[c] SD score is the social desirability score as obtained from the Marlowe-Crowne Desirability Scale.[1]

[d] Body mass index = weight (kg)/height (m)^2.

[e] The nutrient values are in daily amounts specified or they are expressed as percentages.

TABLE 4 *Results of general linear models to assess gender differences in social desirability bias based on pre and post measurements. WATCH Nutrition Validation and Social Desirability Studies, Worcester, MA, USA, 1991–1993*[a]

	Pre measures				Post measures			
	Women		Men		Women		Men	
	b[c]	SE_b[c]	b	SE_b	b	SE_b	b	SE_b[c]
Dependent variable[b]								
Total fat (g)	-3.36	(1.57)	-1.11	(1.79)	-2.06	(0.64)	-0.35	(1.43)
Saturated fatty acids (g)	-0.85	(0.45)	-0.29	(0.67)	-0.62	(0.21)	0.22	(0.49)
Monounsaturated fatty acids (g)	-1.25	(0 56)	-0.48	(0 73)	0.83	(0.28)	-0.27	(0.58)
Polyunsaturated fatty acids (g)	-1.29	(0.51)	-0.32	(0.46)	-0.45	(0.17)	-0.35	(0.39)
Alcohol (g)	-0.17	(0.15)	-0.13	(0.46)	-0.24	(0.28)	-0.09	(0.35)
Total energy (kcal)	-68.05	(30.31)	-38.90	(32.00)	-47.33	(14 35)	-17.31	(32.97)
Non-fat energy (kcal)	-34.83	(16.92)	-20.25	(22.25)	-26.32	(9.33)	13.06	(24.35)
Fat (% energy)	-0.33	(0.23)	-0.13	(0.20)	-0.01	(0.22)	-0.09	(0.21)
Cholesterol (mg)	-7.17	(2.55)	-11.25	(7.56)	-3.68	(2.13)	1.48	(6.00)
Dietary fibre (g)	-0.53	(0.18)	0.05	(0.28)	-0.44	(0.13)	-0.14	(0.39)

[a] All models were fitted with the 24-hour diet recall-derived nutrient variable corresponding to the 7-day diet recall (7DDR)-derived dependent variable as well as social desirability score, body mass index, and interval (in days) between posting and return of the social desirability questionnaire as the independent variables. Models represent stratification according to gender. All results are based on type III (orthogonal) sums of squares.

[b] Models are based on the nutrient score derived from the 7DDR at the beginning of the study (pre measures) and at the end (post measures) as the dependent variables.

[c] b is the regression coefficient and SE_b is its standard error based on fitting the social desirability score as obtained from the Marlowe-Crowne Social Desirability Scale.[1] Significance testing was not conducted because of the exploratory nature of the analysis.

TABLE 5 *Results of general linear models, quartile stratification to assess variation in social desirability bias, according to 24HR-derived fat and total energy scores, WATCH Nutrition and Social Desirability Studies, Worcester, MA, USA, 1991–1993a*

	Pre measures[b]		Post measures[b]	
	b[c]	SE_b[c]	b[c]	SE_b[c]
Fat				
Quartile 1	-1.11	0.85	-1.15	0.64
Quartile 2	0.33	0.65	-0.29	0.32
Quartile 3	-1.42	1.12	-1.48	0.24
Quartile 4	-3.62	2.95	-1.65	1.08
Total energy				
Quartile 1	-18.88	19.56	-22.17	5.76
Quartile 2	6.39	3.42	10.27	4.90
Quartile 3	-22.79	11.22	13.34	24.01
Quartile 4	-72.66	84.96	-13.02	45.27

[a] All models were fitted with the 24-hour diet recall derived nutrient variable corresponding to the 7-day diet recall (7DDR)-derived dependent variable as well as social desirability score, body mass index, and interval (in days) between posting and return of the social desirability questionnaire as the independent variables. Models represent stratification according to 24HR-derived value of the nutrient. All results are based on type III (orthogonal) sums of squares.

[b] Models are based on the nutrient score derived from the 7DDR at the beginning of the study (Pre measures) and at the end (Post measures) as the dependent variables.

[c] b is the regression coefficient and SE_b is its standard error based on fitting the social desirability score as obtained from the Marlowe-Crowne Social Desirability Scale.[1] Significance testing was not conducted because of the exploratory nature of the analysis and very small numbers (n = 10, per group).

investigated in FFQ that are commonly used in epidemiological studies.

The effect of social desirability persisted across most of the data including energy from all non-fat sources and dietary fibre. Several plausible explanations could reconcile the apparent anomaly: statistical analyses examining highly intercorrelated nutrients often produce similar results; among our study subjects there may have been little or no distinction according to nutrient composition of the foods consumed, but more apparent concern with absolute amount of consumption; and the 7DDR was not optimal for quantifying fibre intake.[42] The 7DDR we used in this study focuses mainly on macronutrients. Because the same theoretical constructs apply to the micronutrients and the fact that nutrients tend to be highly intercorrelated, it is likely similar results would be obtained for the micronutrients.

No effect was observed for social approval. Apparently, the desire to avoid criticism around eating behaviour is stronger than is the need to seek approval. It will be important to see if this holds for other populations.

Because social desirability and social approval are considered traits, it was reasonable to administer the forms 2 years after the end of the original study. The administration sequence precluded biasing the dietary results. Delay in returning the questionnaire forms showed results similar to those observed with increased social desirability, but of smaller magnitude. This raises interesting questions about non-response bias.

Possible Consequences of a Social Desirability Bias

The consequences of biased self-report of dietary intake would depend on two general considerations: 1) how social desirability is distributed in the population and its effect across levels of the nutritional variables of interest; and 2) whether social desirability is related to the study endpoint.

If everyone in a study population expressed the same need to provide socially-desirable answers to questions on food intake, its effect would be a uniform reduction in the estimated nutrient score. The bias thus produced would be uniform across the levels of the nutrient distribution. The overall range of the nutrient scores would be identical to the 'true' distribution and the risk estimate of the nutrient-disease relationship would be unaffected. However, the intercept (i.e. the prediction as to where the effect occurs in the nutrient distribution) might be offset significantly.

It is likely that certain individuals have a greater than average need to provide socially desirable answers, as we observed in our study. If the bias associated with social desirability did not differ by level of nutrient intake, it would alter risk estimates across the entire distribution. Were the underestimation to differ by level of nutrient score, as we observed, it would result in biased risk estimates in specific regions of the distribution.

If social desirability were related both to the study endpoint and reporting of dietary data, there is a potential for a classical confounding bias. The effect on the risk estimate could be to increase, attenuate, or reverse it. For follow-up studies, clearly there is no concern with respect to participants' prior knowledge of their disease status. For case-control studies, however, merely knowing one's disease status and having a preconceived notion about a nutrition-disease link sets the stage for confounding. This potential exists irrespective of any true aetiologic relationship between social desirability and the study endpoint.

Though follow-up studies are free from biases related to one's knowledge of a confirmed disease diagnosis, it does not mean that they cannot be classically confounded. For example, with increasing knowledge about psychoneuroimmunology,[47–49] it is becoming clear that a very compliant personality (i.e. a high-scorer on the social desirability scales) may have reduced immunocompetence on parameters such as natural killer (NK) cell activity that may affect directly the probability of developing cancer.[50,51] This also is in keeping with some evidence that a 'cancer-prone' personality scores high on the social desirability scale.[52,53] It is plausible that these same people might tend to under-report high-fat (i.e., socially undesirable) foods. The consequence could be the complete (or even greater) countervailing of a true effect of fat on cancer, thus producing null results or even suggesting a negative effect of these foods when, in fact, they increase risk. Social desirability/approval also may be associated with neuro-endocrine effects[54,55] that may have profound implications for diseases such as breast cancer.

The potential role of social desirability in biasing dietary self-reports should be of major concern in intervention studies where specific recommendations are made to change diet. In these instances, social desirability operates in the context of a potential demand-characteristics bias (i.e. individuals are told to change diet and then to report on dietary intake). In intervention studies, not only is the dietary assessment subject to potential biases because of general population-level messages, but it is being called upon to measure the effect of an explicit intervention aimed at increasing consumption of socially-desirable foods and decreasing consumption of undesirable foods.

Almost without exception, the FFQ currently being used in intervention studies were not intended or designed to measure dietary change as a consequence of an intervention.[56,57]

SUMMARY AND RECOMMENDATION

Diet questionnaires, though ultimately used to produce nutrient scores, are really cognitive and psychosocial test instruments. There are hundreds of published studies that have used the MCSD as part of the validation of a newly-developed measure.[58–60] However, this instrument has not been used in the development of a dietary questionnaire or for adjustment purposes after dietary data are collected. Unlike for some psychological parameters, dietary assessment requires that we do not omit entire food categories, even if it is evident that reported consumption is biased. A more appropriate use of social desirability estimates in this context would be to include the score as a covariate in analysis or as a means of correcting a nutrient score, much as is done with the validity scales of the Minnesota Multiphasic Personality Inventory (MMPI) in psychological practice.

Growing public interest in diet-disease relationships and increasing availability of information on this topic heightens concern about the role of social desirability/approval biases in dietary self-report. That mismeasurement in dietary assessment is a very serious problem is becoming more widely accepted.[61,62] This problem ultimately converges with broad population-based messages to change diet.[63,64] It both begs research into the causes of errors in dietary self-report, confronts us with crucial issues of defining content validity more broadly and identifying comparison criteria in the design of these research studies, and poses an immediate need to quantify and adjust for social desirability in epidemiological studies.

REFERENCES

[1] Crowne D, Marlowe D. A new scale of social desirability independent of psychopathology. *J Consult Clin Psychol* 1960; **24:** 349–54.

[2] Larsen K, Martin H, Ettinger R, Nelson J. Approval seeking, social cost, and aggression: a scale and some dynamics. *J Psychol* 1976; **94:** 3–11.

[3] Edwards A. Social desirability or acquiescence in the MMPI? A case study with the SD scale. *J Abnormal Social Psychol* 1961; **63:** 351–59.

[4] Anastasi A. *Psychological Testing. 6th edn.* New York: MacMillan Publishing Co., 1988.

[5] Edwards A. *The Social Desirability Variable in Personality Assessment and Research.* New York: Dryden, 1957.

[6] Meehl P, Hathaway S. The K factor: a suppressor variable in the MMPI. *J Appl Psychol* 1946; **30:** 525–64.

[7] Marlowe D, Crowne D. Social desirability and responses to perceived situational demands. *J Consult Clin Psychol* 1961; **25:** 109–15.

[8] Larsen K, Martin H, Giles H. Anticipated social cost and interpersonal accommodation. *Hum Commun Res* 1977; **3:** 303–08.

[9] Martin H. A revised measure of approved motivation and its relationship to social desirability. *J Pers Assess* 1984; **48:** 508–16.

[10] Shulman A, Silverman I. Social desirability and need approval: Some paradoxical data and a conceptual reevaluation. *Br J Soc Clin Psychol* 1974; **13:** 27–32.

[11] Evans R. The relationship of the Marlowe-Crowne Scale and its components to defensive preferences. *J Pers Assess* 1979; **43:** 406–10.

[12] Thoreson C, Mahoney M. *Behavioral Self Control.* New York: Holt Rinehart and Winston, 1974.

[13] Goldberg J. Nutrition and health communication: the message and the media over half a century. *Nutr Rev* 1992; **50:** 71–77.

[14] McConaghy J. Adults' belief about the determinants of successful dietary change. *Community Health Studies* 1989; **13:** 492–502.

[15] Roering K, Boush D, Shipp S. *Factors that Shape Eating Patterns: A Consumer Perspective. What is America Eating?* Washington, DC: National Academy Press, 1986, pp. 72–84.

[16] Milio N. Nutrition and health: patterns and policy perspectives in food-rich countries. *Soc Sci Med* 1989; **29:** 413–23.

[17] Drewnowski A. Dietary fats: Perceptions and preferences. *J Am Coll Nutr* 1990; **9:** 431–35.

[18] Suitor C, Gardner J, Willett W. 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.

[19] Bazzarre T, Yuhas J. Comparative evaluation of methods of collecting food intake data for cancer epidemiology studies. *Nutr Cancer* 1983; **5:** 201–14.

[20] Willett W, Sampson L, Stampfer M *et al.* Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 1985; **122:** 51–65.

[21] Larkin F, Metzner H, Thompson F, Flegal K, Guire K. Comparison of estimated nutrient intakes by food frequency and dietary records in adults. *J Am Diet Assoc* 1989; **89:** 215–23

[22] Worsley A, Baghurst K, Leitch D. Social desirability response bias and dietary inventory responses. *Human Nutr Appl Nutr* 1984; **38A:** 29–35.

[23] Southgate D. Obese deceivers. *Br Med J* 1986; **292:** 1692–93.

[24] Kim W, Mertz W, Judd J, Marshall M, Kelsay J, Pather E. Effect of making duplicate food collections on nutrient intakes calculated from diet records. *Am J Clin Nutr* 1984; **40:** 1333–37.

[25] Jain M, Howe G, Harrison L, Miller A. A study of repeatability of dietary data over a seven-year period. *Am J Epidemiol* 1989; **129:** 422–29.

[26] Colditz G, Willett W, Stampfer M *et al.* The influence of age, relative weight, smoking, and alcohol intake on the reproducibility of a dietary questionnaire. *Int J Epidemiol* 1987; **16:** 392–98.

[27] Block G, Dresser C, Hartman A, Carroll M. Nutrient sources in the American diet: Quantitative data from the NHANES II survey: II. macronutrients and fats. *Am J Epidemiol* 1985; **122:** 27–39.

[28] Wardle J, Beales S. Restraint, body image and food attitudes in children from 12 to 19 years. *Appetite* 1986; **7:** 209–17.
[29] Polivy J, Herman C, Hackett R, Kuleshnvk I. The effects of self-attention and public attention on eating in restrained and unrestrained subjects. *J Person Soc Psychol* 1986; **6:** 1253–60.
[30] Herman C, Polivy J, Silver R. Effects of an observer on eating behavior. The induction of 'sensible' eating. *J Personality* 1979; **47:** 85–99.
[31] Bancroft J, Cook A, Williamson L. Food craving, mood and the menstrual cycle. *Psychol Med* 1988; **18:** 855–60.
[32] 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.
[33] Wurtman J, Wurtman R, Growdon J, Henry P, Lipscomb A, Zeisel S. Carbohydrate craving in obese people: suppression by treatments affecting serotoniergic transmission. *Int J Eating Disorders* 1981; **1:** 2–15.
[34] Prentice A, Black A, Coward W *et al.* High levels of energy expenditure in obese women. *Br Med J* 1986; **292:** 983–87.
[35] Bandini L, Schoeller D, Cyr H, Dietz W. Validity of reported energy intake in obese and nonobese adolescents. *Am J Clin Nutr* 1990; **52:** 421–25.
[36] Johnson W, Corrigan S, Schlundt D, Dubbert P. Dietary restraint and eating behavior in the natural environment. *Addic Behav* 1990; **15:** 285–90
[37] Horm J, Anderson K. Who in America is trying to lose weight?: National Health Interview Survey. In: Proceedings of Methods for Voluntary Weight Loss and Control: An NIH Technology Assessment Conference, 1992, pp. 50–57.
[38] Greeno C, Wing R. Stress-induced eating. *Psychological Bull* 1994; **115:** 444–64.
[39] Klem M, Klesges R, Bene C, Mellon M. A psychometric study of restraint: the impact of race, gender, weight, and marital status. *Addictive Behav* 1990; **15:** 147–51.
[40] Seim H, Fiola J. A comparison of attitudes and behaviors of men and women toward food and dieting. *Fam Pract Res J* 1990; **10:** 57–63.
[41] Ganong W. Neural basis on instinctual behavior and emotions. *Review of Medical Physiology.* Norwalk, CT: Appleton & Lange, 1991, pp. 237–47.
[42] Hebert J, Ockene I, Botelho L, Luippold R, Merriam P, Saperia G. Development and validation of a seven-day diet recall. In: American Public Health Association 120th Annual Meeting. Washington, DC, 1992.
[43] Posner B, Martin-Munley S, Smigelski C *et al.* Comparison of techniques for estimating nutrient intake: The Framingham Study. *Epidemiology* 1992; **3:** 171–77.
[44] Hebert J, Miller D, Barone J, Engle A. Assessment of change in fat intake in an intervention study. In: *American Public Health Association – 119th Annual Meeting*. Atlanta, GA: APHA, 1991.
[45] Miller D, Hebert J, Barone J, Engle A. Comparison of dietary assessment methods: advantages of random 24 hour recalls. In: *American Public Health Association – 119th Annual Meeting*. Atlanta, GA: APHA, 1991.
[46] SAS. *SAS Users Guide – Statistics Version 6.5.* Cary, NC: SAS Institute, 1993.
[47] Glaser R, Kiecolt-Glaser J, Stout J, Tar K, Speicher C, Holliday J. Stress-related impairments in cellular immunity. *Psych Res* 1985; **16:** 233–39.
[48] Fox B, Temoshok L, Dreher H. Mind-body and behavior in cancer incidence. *Advances* 1988; **5:** 41–56.
[49] Fox B. Psychogenic factors in cancer, especially its incidence. In: Maes S, Spielberger C, Defares P, Sarason I, (eds). *Topics in Health Psychology.* New York: John Wiley, 1988, pp. 37–55.
[50] Levy S, Herberman R, Lippman M, d'Angelo T. Correlation of stress factors with sustained depression of natural killer cell activity and predicted prognosis in patients with breast cancer. *J Clin Oncol* 1987; **5:** 348–53.
[51] Hebert J, Barone J, Reddy M, Backlund J. Natural killer cell activity in a dietary fat intervention trial. *Clin Immunol Immunopathol* 1990; **54:** 103–16.
[52] Kune G, Kune S, Watson L, Bahnson C. Personality as a risk factor in large bowel cancer: data from the Melbourne Colorectal Cancer Study. *Psychological Med* 1991; **21:** 29–41.
[53] Eysenck H. Personality, stress and cancer: Prediction and prophylaxis. *Br J Med Psychol* 1988; **61:** 57–75.
[54] Hauenstein E. Young women and depression. Origin, outcome, and nursing care. *Nurs Clin North Am* 1991; **26:** 601–12.
[55] Daruna J, Morgan J. Psychosocial effects on immune function: neuroendocrine pathways. *Psychosomatics* 1990; **31:** 4–12.
[56] Hebert J, Harris D, Sorensen G, Stoddard A, Hunt M, Morris D. A worksite nutrition intervention: its effects on the consumption of cancer-related nutrients. *Am J Public Health* 1993; **83:** 391–94.
[57] Abrams D, Boutwell W, Grizzle J, Heimendinger J, Sorensen G, Varnes J. Cancer control at the workplace: The Working Well Trial. *Prev Med* 1994; **23:** 15–27.
[58] Maher B. A reader's, writer's and reviewer's guide to evaluation of clinical research reports. *J Consult Clin Psychol* 1978; **46:** 835–38.
[59] Mischel W. *Personality & Assessment.* New York: John Wiley, 1968.
[60] Crowne D. *The Experimental Study of Personality.* Hillsdale N J: Erlbaum, 1979.
[61] Hebert J, Miller D. Methodologic considerations for investigating the diet-cancer link. *Am J Clin Nutr* 1988; **47:** 1068–77.
[62] Freudenheim J, Marshall J. The problem of profound mismeasurement and the power of epidemiological studies of diet and cancer. *Nutr Cancer* 1988; **11:** 243–50.
[63] US Department of Health and Human Services. *The Surgeon General's Report on Nutrition and Health.* DHHS (PHS Publ. No. 88-50210). Washington DC: US GPO, 1988.
[64] National Research Council. *Diet and Health: Implications for Reducing Chronic Disease Risk.* Washington, DC: National Academy Press, 1989.

(Revised version received October 1994)