A Search for Recall Bias in a Case-Control Study of Diet and Breast Cancer

L HOLMBERG,* E M OHLANDER,† T BYERS,‡ M ZACK,§ A WOLK,** A BRUCE,† R BERGSTRÖM,** L BERGKVIST³ AND H O ADAMI* 


Background. In retrospective studies of dietary habits and breast cancer risk, recall bias is a concern since diet has been publicized as a cause of breast cancer.

Methods. In a case-control study of diet and breast cancer risk nested within a cohort of women screened with mammography, we contrasted answers to a retrospective dietary interview with answers to a dietary questionnaire which was filled out before any diagnostic procedures for breast cancer were undertaken. The source population was all women aged 40–74 in two counties in Sweden invited to mammographic screening and asked to fill out a questionnaire before the screening. Cases and controls were subsequently defined—matched on age, county of residence, and time of mammography — and approached for an interview.

Results. In all, 265 cases and 431 controls participated in the study. Means of monthly frequencies differed between the questionnaire and the interview for both cases and controls to a similar degree in all food groups. The percentage of agreement in the questionnaire and the interview's classifications of study subjects into quartiles of monthly intake varied between 31% and 57%. Kappa statistics in all food groups were below 0.41. These measures of agreement did not differ between cases and controls. The confidence intervals for odds ratios of breast cancer risk obtained from the two measurements overlapped for all food groups. In a regression analysis, case subjects with low responses on the questionnaire about intake of meat, snacks, and coffee and tea gave higher responses on interview than did controls who had low questionnaire responses for these food groups. The reverse was also true: cases' responses that were high on the questionnaire were lower on interview for these food groups than were controls' responses.

Conclusions. We found few signs of recall bias, and the few indications of a differential misclassification that we found were not in food groups that have been publicly discussed as causes of breast cancer.

Keywords. Breast cancer, diet, case-control study, mammography

Case-control studies are susceptible to recall bias because study subjects might recall a prior exposure differently depending on their disease status.¹ Recall bias is a particular concern when an exposure that can be determined only by self-report has been publicized as a cause for the disease under study. Because dietary habits have been publicly discussed in most countries with high breast cancer incidence as affecting breast cancer risk, the association between dietary fats and breast cancer observed in many case-control studies² could be viewed as resulting from recall bias.³

Although many studies have examined the validity of food frequency dietary questionnaires (the method commonly used in case-control studies), only five studies have analysed recall bias in relation to dietary history obtained after a breast cancer diagnosis.²-⁵ All but one³ suggested that this bias was minor or negligible. Importantly, however, no study has exactly reproduced the way in which cases and controls are approached in an incident case-control study. Some of the investigators used cases identified during a 5- to 7-year period,⁴,⁵,⁷ and the questionnaire administered after case status was established was sent out one year or more after diagnosis. Hislop et al.⁶ contrasted questionnaires administered at different time intervals after cases and controls were ascertained. In one study, the
investigators asked the cases and controls about their diets of 3 years earlier. This paper reports our findings from a case-control study of diet breast cancer risk, in which the design allowed us to determine whether incident breast cancer cases systematically differed from controls in how they retrospectively reported their dietary habits in the 6-month period before diagnosis. The question of recall bias was addressed by contrasting the answers to a dietary interview in a case-control study with the answers to a dietary questionnaire that had been completed by all study subjects before breast cancer screening.

SUBJECTS AND METHODS

Study Design
A case-control study to analyse relationships between dietary history and breast cancer risk was planned when the first rounds of population-based public health care mammography screening started in two counties in central Sweden in 1986. The source population was all women aged 40–74 years living in one county (Uppsala) and those aged 40–70 in another county (Västmanland) who had not had breast cancer diagnosed before screening. The screening interval was 18 months for women 40–55 and 24 months for older targeted women. A food frequency questionnaire was mailed to each woman along with the invitation to participate in the breast cancer screening. The questionnaire was to be returned to the screening centre before the screening examination. The participation rate in the screening programme was 86% and 89% in the two counties, respectively. Of the women who came to the screening, 81% had returned the questionnaire implying that about 71% of the source population was included in the study cohort. All women included were screened from March 1987 through December 1990, and they had completed a dietary questionnaire before the screening.

Cases
Eligible as cases in this analysis were all women from the source population who had completed the dietary questionnaire, had participated in the first screening round, and had subsequently been diagnosed with invasive breast cancer. The cancer could have been diagnosed at the first screening or in the following 2 years (Table 1). Cases were identified both from pathology reports and by the screening centres.

Controls
Eligible as controls were those women who had completed the questionnaire and had participated in the first screening round. Controls were selected by age group-matching based on the numbers of breast cancer cases expected in each 5-year age group (40–74) on the basis of a previous screening trial per county per month, given the anticipated number of women screened per month. The planned case-control ratio was 1:1. The number of controls exceeded the number of cancer cases, however, because the cancer detection rates at screening did not reach the expected levels.

Questionnaire
The first part of the questionnaire included questions about age, marital status, education, parity, occurrence of breast cancer in first-degree relatives (mother, sister, daughter), weight, and height. Other questions followed about usual type of diet (e.g. vegetarian), type and amount of fat used on bread, type and amount of milk consumed, and type of fat used in food preparation. The food frequency part of the questionnaire included 60 foods and food groups (Appendix). Respondents were asked to report their usual intake of each food during the past 6 months. For each food item, eight response categories were defined: never/seldom, 1–3 times a month, once a week, 2–3 times a week, 4–6 times a month, once a week, 2–3 times a week, and ≥4 times a day.

The questionnaire has been validated against 4 × 7 days of prospective dietary records completed by 129

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Participation rate and demographic characteristics of cases and controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
</tr>
<tr>
<td>Total number</td>
<td>380</td>
</tr>
<tr>
<td>Number interviewed (%)</td>
<td>276 (73%)</td>
</tr>
<tr>
<td>Participants in study:</td>
<td></td>
</tr>
<tr>
<td>Total number included</td>
<td>265</td>
</tr>
<tr>
<td>Median time elapsed between questionnaire and interview (months)</td>
<td>7</td>
</tr>
<tr>
<td>Mean age (SD)</td>
<td>59.0 (9.2)</td>
</tr>
<tr>
<td>No. of children</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>33 (12%)</td>
</tr>
<tr>
<td>1</td>
<td>66 (25%)</td>
</tr>
<tr>
<td>2</td>
<td>89 (34%)</td>
</tr>
<tr>
<td>3–4</td>
<td>69 (26%)</td>
</tr>
<tr>
<td>≥5</td>
<td>8 (3%)</td>
</tr>
<tr>
<td>Body mass index</td>
<td></td>
</tr>
<tr>
<td>&lt;23.0</td>
<td>63 (24%)</td>
</tr>
<tr>
<td>23.0–24.9</td>
<td>71 (27%)</td>
</tr>
<tr>
<td>25.0–27.4</td>
<td>65 (25%)</td>
</tr>
<tr>
<td>≥27.5</td>
<td>62 (23%)</td>
</tr>
<tr>
<td>No special diet</td>
<td>260 (98%)</td>
</tr>
<tr>
<td>Breast cancer in first-degree relative</td>
<td>28 (11%)</td>
</tr>
<tr>
<td>More than 9 years of primary school</td>
<td>44 (17%)</td>
</tr>
</tbody>
</table>

* No. women with <30 missing answers in the questionnaire.
women (Wolk A, personal communication). When age-
dependent portion sizes were used, Pearson correlation
coefficients from 0.35 to 0.82 were obtained for calcu-
lations of intake of different nutrients. The coefficients
were 0.47 for total fats, 0.53 for dietary fibre, 0.39 for
carotene and 0.82 for alcohol.

Interview
While the food frequency questionnaire was designed to
obtain information from a very large cohort of
women and could be used as an assessment of dietary
intake 6 months before the diagnosis of breast cancer,
an interview study was designed to obtain information
on diet from 15 years of age until present time. The
interview approach was used since we were specifically
interested in risk of breast cancer related to age of
exposure, latency and recency and average lifetime
exposure.

If a woman was eligible as a case or was selected
as a control, she first received a letter with information
about the study. She was then called on the tele-
phone for her consent and to make an appointment for
a later telephone interview. Interviews took place about
6 months after screening (when primary treatment of
the cancer had been completed for cases, or after a
comparable time for controls).

At the beginning of the telephone interview, ‘dietary
periods’ for each women were identified. A dietary
period was defined as a succession of years for which a
woman stated she had not changed her eating habits.
Dietary periods usually began with life events such as
marriage and childbirth or with a change in views about
‘healthy’ eating habits. The interview always proceeded
from the earliest identified period to the latest, i.e. the
6 months preceding breast cancer screening. In this
analysis, we used only information from the latest
dietary period. For each dietary period, the respondents
were asked about consumption frequencies for the same
dietary factors, <1% of the study subjects had missing
information.

Each nutritionist interviewed a similar proportion of
cases and controls.

Missing Values and Exclusions
In the dietary questionnaire, 72% of the study subjects
had fewer than five missing values concerning the diet-
ary information, and 86% had fewer than 10 missing
values. Missing values were most frequently in the food
groups in which several alternatives with closely
related items were given, in particular in the milk,
butter and margarine, and bread groups. Interviews in a
sample of women revealed that they often had not filled
out one of the alternatives when they never or seldom
ate that particular item.10 Missing values in the ques-
tionnaire also strongly coincided with responses of
‘never’, ‘seldom’, or ‘<4 times per month’ in the inter-
view. Missing values were therefore interpreted as
‘never/seldom’. Women with ≥30 items missing in the
questionnaire (4.4% of all those interviewed) were
excluded from this analysis (Table 1). For some of the
analyses (as stated below) only women with complete
questionnaires were analysed. There were no missing
values in the interviews.

Data on education, height, or weight were missing for <2% of both cases and controls. For other non-
dietary factors, <1% of the study subjects had missing
information.

Analyses
The responses to the questionnaire and the interview
were converted into frequencies per month. We classi-
ﬁed foods into nine groups: dairy products, vegetables,
fruits, meat (including fish and egg dishes), cereals,
snacks (including popcorn, potato chips, buns, ice
cream, cookies, etc.), coffee and tea, alcoholic bever-
ages, and a group of foods we believe are perceived in
Sweden to be fatty and unhealthy (Appendix).

We compared the mean monthly frequencies for food
groups in the questionnaire and the interview with
paired t-tests for cases and controls separately. The mean
differences between responses to the questionnaire and
the interview for cases and controls were compared
with unpaired t-tests. We formed quartiles of monthly
intake for each of the nine food groups from both the
questionnaire and interview responses. The quartile
boundaries were deﬁned by the response of the con-
trols. We constructed sets of 4 × 4 tables comparing
how participants were grouped in quartiles depending
on questionnaire or interview responses, and then cal-
culated percentage of agreement and kappa statistics.11,12

We regressed the interview response in frequencies
per month on the questionnaire response. Thus model
construction enabled us to test whether the relationship
### Table 2 Means of monthly frequencies of intake of different food groups in questionnaire (Q) and interview (I) for cases and controls

<table>
<thead>
<tr>
<th>Food Group</th>
<th>Cases Q</th>
<th>Cases I</th>
<th>% mean difference: Q - I</th>
<th>% mean difference: Q - I</th>
<th>Controls Q</th>
<th>Controls I</th>
<th>% mean difference: Q - I</th>
<th>% mean difference: Q - I</th>
<th>p*</th>
<th>p#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dairy products</td>
<td>157.8</td>
<td>168.6</td>
<td>-6.8*</td>
<td>(-5.9)</td>
<td>147.9</td>
<td>164.7</td>
<td>-11.4*</td>
<td>(-15.7*)</td>
<td>0.34</td>
<td>0.31</td>
</tr>
<tr>
<td>Vegetables</td>
<td>79.0</td>
<td>78.6</td>
<td>0.5</td>
<td>(2.6)</td>
<td>79.2</td>
<td>83.3</td>
<td>-5.2</td>
<td>(-4.6)</td>
<td>0.21</td>
<td>0.25</td>
</tr>
<tr>
<td>Fruits</td>
<td>55.3</td>
<td>49.2</td>
<td>11.0*</td>
<td>(8.2)</td>
<td>52.7</td>
<td>49.4</td>
<td>6.3</td>
<td>(3.1)</td>
<td>0.35</td>
<td>0.36</td>
</tr>
<tr>
<td>Meat</td>
<td>50.0</td>
<td>54.4</td>
<td>-8.8*</td>
<td>(-2.9)</td>
<td>49.2</td>
<td>55.7</td>
<td>-13.2*</td>
<td>(-6.3*)</td>
<td>0.34</td>
<td>0.29</td>
</tr>
<tr>
<td>Cereals</td>
<td>111.8</td>
<td>120.8</td>
<td>-8.1*</td>
<td>(-7.8)</td>
<td>109.7</td>
<td>118.8</td>
<td>-8.3*</td>
<td>(-2.9)</td>
<td>0.99</td>
<td>0.52</td>
</tr>
<tr>
<td>Snacks</td>
<td>62.7</td>
<td>80.6</td>
<td>-28.6*</td>
<td>(-18.6*)</td>
<td>61.0</td>
<td>79.2</td>
<td>-29.8*</td>
<td>(-11.9)</td>
<td>0.94</td>
<td>0.36</td>
</tr>
<tr>
<td>Coffee, tea</td>
<td>85.7</td>
<td>92.1</td>
<td>-7.5</td>
<td>(-9.0)</td>
<td>84.3</td>
<td>93.6</td>
<td>-11.0*</td>
<td>(-7.5*)</td>
<td>0.35</td>
<td>0.76</td>
</tr>
<tr>
<td>Alcoholic beverages</td>
<td>6.8</td>
<td>4.2</td>
<td>8.8</td>
<td>(1.5)</td>
<td>6.7</td>
<td>5.5</td>
<td>17.9</td>
<td>(1.3)</td>
<td>0.48</td>
<td>0.88</td>
</tr>
<tr>
<td>'Fatty foods'</td>
<td>131.9</td>
<td>131.6</td>
<td>0.2</td>
<td>(-3.0)</td>
<td>123.2</td>
<td>126.0</td>
<td>-2.3</td>
<td>(-1.4)</td>
<td>0.59</td>
<td>0.86</td>
</tr>
</tbody>
</table>

* t-test contrasting case and controls.

b Values for analysis only on cases and controls with complete questionnaire.

* P < 0.05, paired t-test.

---

between questionnaire and interview values differed between cases and controls. In all models, we accommodated the study design by including county of residence, age at screening, and the date of screening as independent variables. In some models, we also included as independent variables possible confounders (age at first birth, number of children, marital status, level of education, family history of breast cancer, body mass index, and type of diet), categorized as indicator variables. We used a log transformation of the monthly food group frequencies. Body mass index was calculated as weight (kg)/height (m)^2. The regression models and the analyses of agreement were repeated only including women without missing values in the questionnaire.

Breast cancer risk was modelled with logistic regression. All models included dummy variables that represented the quartiles of intake of the food group under study and the stratification variables. The sum of the absolute values of the difference between a participant’s questionnaire response and interview response for monthly intake for each food group was used to estimate agreement between the two kinds of dietary measurement. For most food groups, most women had values 10 through 20, but some women had much larger differences. The two quartiles of these values closest to zero were judged to have a good agreement, the third, to have a fair agreement, and the fourth, to have a poor agreement. Differences in the odds ratios obtained in logistic regression analyses stratified according to agreement were tested with indicator terms for these three subgroups and with their interaction terms for the food group in question. To find out if odds ratios obtained from questionnaire and interview data differed from one another, we tested if the parameter estimates for the difference between the questionnaire and interview, adjusted for the stratification variables, were statistically significant in a logistic regression model. The level of statistical significance in all tests was considered to be α = 0.05.

**RESULTS**

Except for a higher proportion among cases of women with a family history of breast cancer (11%, cases; 6%, controls) and a trend for controls to have a higher parity, other basic demographic characteristics did not differ between cases and controls (Table 1).

For both cases and controls, the mean monthly frequencies in the questionnaire differed significantly from, and were usually less than, those in the interview for all food groups except vegetables, alcoholic beverages, and fatty foods (Table 2). For cases, but not for controls, the mean monthly frequency for fruits differed significantly between the questionnaire and the interview, being substantially higher in the questionnaire than in the interview as opposed to the pattern in most food groups. Among both cases and controls, the standard deviations of the interview responses were smaller than those of the questionnaire responses for all
food groups except alcoholic beverages, for which they were similar. The mean differences between the questionnaire and the interview responses as a percentage of the questionnaire response are similar for cases and controls (Table 2). The analyses were also repeated for cases and controls that had a complete questionnaire. Overall a somewhat better agreement between questionnaire and interview response could be seen, but the overall pattern with higher interview responses and an absence of a statistically significant difference between cases and controls was not altered (Table 2).

Except for alcoholic beverages, the percentage of agreement between the questionnaire and the interview was low overall (Table 3). The kappa statistics were statistically significant for all food groups and indicated some relation between the questionnaire and the interview response. However, the fact that these estimates—except for alcoholic beverages—were all less than 0.3 indicated minimal agreement. Only the kappa statistics for fatty foods showed a statistically significantly better agreement between questionnaire and interview responses among cases than among controls. Also for this set of analyses, we repeated the calculations on the subset of cases and controls that had a complete response to the questionnaire. As in Table 2, a tendency for a higher degree of agreement between the questionnaire and interview could be seen, but the contrast between cases and controls was not altered (Table 3).

We regressed the log of the interview response in the basic model on the log of the questionnaire response (Table 4). We took into account the case-control status and the design variables of the study. A basic model (number 1 in Table 4) was extended with an interaction term (between case-status and log of questionnaire-response) to explore whether any effect modification by case status existed (model 2) and finally, with the covariates age at first birth, parity, marital status, education, family history, and type of diet (model 3), to see whether any of these confounded the relationship.

In the ideal situation (exact agreement between questionnaire and interview responses and no difference between cases and controls), the regression lines would both pass through the origin and have a slope equal to 1.0 (Figure 1a). In this study, however, the slopes are shallower for all food groups. Unless this finding is only a regression fallacy due to measurements with large random errors, it implies that the interview responses underestimate the questionnaire responses in the upper ranges. Since the intercepts, except that for meat, do not differ statistically from zero, the curves do cross the y-axis close to zero. Thus, the women who responded on the questionnaire that they seldom or never ate foods from a specific food group gave the same response during the interview. For meat, the statistically significant and positive intercept in model 1 indicates that the women who responded on the questionnaire that they never or seldom ate meat later reported that they had eaten meat (Figure 1b). Still, for meat in model 1, no statistically significant difference by case-control status emerged; the case term (0.02 units) being within the limits of random variation.

In model 1 (Table 4), the only significant difference by case-control status appears for snacks (Figure 1c).
TABLE 4 Regression of log interview response (frequency intake per month) on the log of the questionnaire response (log of Q) and taking case-control status into account. Table shows coefficients (SE). All models adjusted for age, county, and date of mammography. Second model also including an interaction term between case status and log of Q. Third model in addition to model two adjusted for age at first birth, number of children, marital status, education, family history, and type of diet

<table>
<thead>
<tr>
<th>Dairy</th>
<th>Vegetables</th>
<th>Fruit</th>
<th>Meat</th>
<th>Cereals</th>
<th>Snacks</th>
<th>Coffee, tea</th>
<th>Alcohol</th>
<th>'Fatty foods'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Intercept</td>
<td>2.7 (2.6)</td>
<td>4.2 (2.2)</td>
<td>-1.3 (3.1)</td>
<td>8.3 (2.0)*</td>
<td>2.8 (2.1)</td>
<td>0.90 (4.6)</td>
<td>-2.5 (2.3)</td>
<td>-3.7 (6.3)</td>
</tr>
<tr>
<td>Log Q^*</td>
<td>0.29 (0.03)</td>
<td>0.26 (0.04)</td>
<td>0.40 (0.03)</td>
<td>0.35 (0.03)</td>
<td>0.19 (0.03)</td>
<td>0.52 (0.03)</td>
<td>0.40 (0.03)</td>
<td>0.63 (0.03)</td>
</tr>
<tr>
<td>case</td>
<td>-0.004 (0.04)</td>
<td>-0.34 (0.32)</td>
<td>-0.004 (0.04)</td>
<td>-0.02 (0.03)</td>
<td>0.02 (0.03)</td>
<td>0.12 (0.06)*</td>
<td>0.02 (0.03)</td>
<td>0.04 (0.09)</td>
</tr>
<tr>
<td>2) Intercept</td>
<td>2.9 (2.6)</td>
<td>4.2 (2.2)</td>
<td>-1.2 (3.1)</td>
<td>8.3 (2.0)*</td>
<td>2.9 (2.1)</td>
<td>0.60 (4.6)</td>
<td>-2.9 (2.3)</td>
<td>-3.9 (6.3)</td>
</tr>
<tr>
<td>Log Q^*</td>
<td>0.26 (0.04)</td>
<td>0.31 (0.04)</td>
<td>0.39 (0.04)</td>
<td>0.42 (0.04)</td>
<td>0.17 (0.03)</td>
<td>0.59 (0.04)</td>
<td>0.49 (0.04)</td>
<td>0.61 (0.06)</td>
</tr>
<tr>
<td>case</td>
<td>-0.32 (0.32)</td>
<td>-0.09 (0.26)</td>
<td>-0.12 (0.21)</td>
<td>0.61 (0.22)</td>
<td>-0.28 (0.28)</td>
<td>0.82 (0.25)*</td>
<td>0.99 (0.26)*</td>
<td>-0.07 (0.21)</td>
</tr>
<tr>
<td>case*logQ</td>
<td>-0.07 (0.07)</td>
<td>0.01 (0.06)</td>
<td>0.03 (0.05)</td>
<td>-0.17 (0.06)*</td>
<td>0.06 (0.06)</td>
<td>-0.19 (0.06)*</td>
<td>-0.23 (0.06)*</td>
<td>0.06 (0.09)</td>
</tr>
<tr>
<td>3) Intercept</td>
<td>2.7 (2.7)</td>
<td>5.5 (2.3)*</td>
<td>-2.1 (3.2)</td>
<td>8.1 (2.1)*</td>
<td>2.6 (2.2)</td>
<td>-1.9 (4.7)</td>
<td>-3.9 (2.4)</td>
<td>-4.3 (6.6)</td>
</tr>
<tr>
<td>Log Q^*</td>
<td>0.25 (0.04)</td>
<td>0.30 (0.04)</td>
<td>0.39 (0.04)</td>
<td>0.42 (0.04)</td>
<td>0.17 (0.04)</td>
<td>0.59 (0.04)</td>
<td>0.49 (0.04)</td>
<td>0.63 (0.06)</td>
</tr>
<tr>
<td>case</td>
<td>-0.32 (0.32)</td>
<td>-0.08 (0.26)</td>
<td>-0.10 (0.21)</td>
<td>0.67 (0.22)</td>
<td>-0.30 (0.28)</td>
<td>0.80 (0.25)*</td>
<td>1.05 (0.27)*</td>
<td>-0.04 (0.21)</td>
</tr>
<tr>
<td>case*logQ</td>
<td>0.07 (0.06)</td>
<td>0.01 (0.06)</td>
<td>0.02 (0.05)</td>
<td>-0.18 (0.06)*</td>
<td>-0.07 (0.06)</td>
<td>-0.18 (0.07)*</td>
<td>-0.23 (0.06)*</td>
<td>0.04 (0.09)</td>
</tr>
</tbody>
</table>

* All of Log Q estimates are statistically significant (P < 0.05).
* P < 0.05.

The intercept is 0.90 for controls but 1.02 (= 0.90 + 0.12) for cases. Thus, cases who reported on the questionnaire that they never or seldom ate snacks later reported that they had eaten more snacks than comparable controls.

Model 1 does not allow the slopes for the regression lines to vary by case-control status within a food group, so the slopes for the regression lines for cases and controls are the same, as, for example, for snacks (Figure 1c). Model 2 (Table 4) allows the slopes to vary. For example, for snacks the interaction term is statistically significant and negative, and indicates a shallower slope for the cases than for the controls, but the cases still have a higher intercept (Figure 1d). This same type of pattern also occurs with meat and coffee or tea. The addition of other covariates to the model (model 3, Table 4) does not change the patterns appreciably.
**Table 5** Odds ratios (OR) for breast cancer risk by food group in highest quartile of intake as compared with the lowest based on questionnaire (Q) and interview (I) data, respectively; study subjects also subdivided according to agreement between Q and I. 95% confidence limits (CI) are given for the total sample and for statistically significant findings in the subgroups.

<table>
<thead>
<tr>
<th>Food Group</th>
<th>All OR (95% CI)</th>
<th>Good agreement Q/I OR</th>
<th>Not good agreement Q/I OR</th>
<th>Bad agreement Q/I OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dairy products</td>
<td>Q 1.3 (0.9-2.0)</td>
<td>2.5* (1.3-4.9)</td>
<td>0.8 (0.5-1.3)</td>
<td>0.9 (0.6-1.3)</td>
</tr>
<tr>
<td></td>
<td>I 1.2 (0.8-1.8)</td>
<td>1.8 (1.2-2.5)</td>
<td>1.2 (0.8-1.6)</td>
<td>0.6 (0.4-0.9)</td>
</tr>
<tr>
<td>Vegetables</td>
<td>Q 1.0 (0.6-1.5)</td>
<td>0.8 (0.6-1.3)</td>
<td>1.1 (0.8-1.5)</td>
<td>1.3 (0.9-1.7)</td>
</tr>
<tr>
<td></td>
<td>I 0.6 (0.4-1.0)</td>
<td>0.4b (0.3-0.6)</td>
<td>0.6 (0.4-0.8)</td>
<td>1.0 (0.7-1.3)</td>
</tr>
<tr>
<td>Fruits</td>
<td>Q 1.0 (0.6-1.5)</td>
<td>0.9 (0.7-1.1)</td>
<td>0.7 (0.5-1.0)</td>
<td>1.2 (0.9-1.5)</td>
</tr>
<tr>
<td></td>
<td>I 1.4 (0.9-2.3)</td>
<td>1.3 (1.1-1.6)</td>
<td>1.0 (0.8-1.2)</td>
<td>2.5 (1.7-3.3)</td>
</tr>
<tr>
<td>Meat</td>
<td>Q 1.1 (0.7-1.7)</td>
<td>1.3 (1.1-1.7)</td>
<td>1.4 (1.2-1.7)</td>
<td>0.6 (0.4-0.8)</td>
</tr>
<tr>
<td></td>
<td>I 0.8 (0.5-1.2)</td>
<td>0.6 (0.4-0.9)</td>
<td>0.6 (0.4-0.9)</td>
<td>1.3 (0.9-1.7)</td>
</tr>
<tr>
<td>Cereals</td>
<td>Q 1.0 (0.7-1.6)</td>
<td>1.2 (1.0-1.4)</td>
<td>0.8 (0.6-1.0)</td>
<td>1.0 (0.8-1.2)</td>
</tr>
<tr>
<td></td>
<td>I 1.0 (0.6-1.5)</td>
<td>1.0 (0.8-1.2)</td>
<td>1.2 (1.0-1.5)</td>
<td>0.7 (0.5-0.9)</td>
</tr>
<tr>
<td>Snacks</td>
<td>Q 1.0 (0.7-1.6)</td>
<td>1.1 (0.9-1.3)</td>
<td>2.3 (2.0-2.6)</td>
<td>0.5 (0.3-0.7)</td>
</tr>
<tr>
<td></td>
<td>I 1.2 (0.8-1.9)</td>
<td>1.3 (1.1-1.5)</td>
<td>1.9 (1.7-2.2)</td>
<td>0.8 (0.6-1.0)</td>
</tr>
<tr>
<td>Coffee, tea</td>
<td>Q 1.1 (0.7-1.9)</td>
<td>1.4 (1.2-1.7)</td>
<td>1.0 (0.8-1.2)</td>
<td>0.8 (0.6-1.0)</td>
</tr>
<tr>
<td></td>
<td>I 1.0 (0.6-1.6)</td>
<td>1.2 (1.0-1.5)</td>
<td>0.8 (0.6-1.0)</td>
<td>0.8 (0.6-1.0)</td>
</tr>
<tr>
<td>Alcoholic beverages</td>
<td>Q 1.4 (1.0-2.1)</td>
<td>1.3 (1.1-1.6)</td>
<td>1.3 (1.1-1.6)</td>
<td>2.5* (2.0-3.0)</td>
</tr>
<tr>
<td></td>
<td>I 1.3 (0.9-2.0)</td>
<td>1.5 (1.2-2.0)</td>
<td>0.5 (0.3-0.7)</td>
<td>2.5* (2.0-3.0)</td>
</tr>
</tbody>
</table>

* 95% CI: 1.3-4.9
b 95% CI: 0.2-0.8.
c 95% CI: 1.1-5.6.
d 95% CI: 1.1-5.3.

We determined if risk of breast cancer by food group differed by method of gathering information (questionnaire or interview) and by degree of agreement between the two methods (Table 5). In the whole data set, the differences in odds ratios obtained in the questionnaire and in the interview were numerically largest for meat, fruits, and vegetables. The difference appears most striking for vegetables, since the interview data suggest an almost statistically significant protective effect, but the questionnaire data show no such effect. For three food groups (dairy products, vegetables, and alcoholic beverages), a gradient in the odds ratios for both questionnaire and interview occurs over the subgroups with different agreement (Table 5). However, recall bias did not appear to be dependent on the degree of agreement: i.e. the questionnaire versus interview difference was not larger in any of the subgroups than expected from looking at the whole data set. No statistically significant differences were found between the odds ratios when they were compared with respect to data source (questionnaire or interview) or to subgrouping according to degree of agreement.

**DISCUSSION**

We wanted to see whether women with a newly diagnosed breast cancer might have reported their dietary habits differently than did women without breast cancer when they were interviewed about 6 months later. During the period of data collection, intake of vegetables, fatty foods and alcohol had been discussed as related to breast cancer risk and theoretically one would have suspected that any signs of recall bias would have been largest for these three food groups. In this analysis, we have assumed that the questionnaire, completed before mammographic screening, provided an unbiased estimate of the study subjects' dietary intake during the 6 months before screening. Although questionnaire responses differed substantially from interview responses in both cases and controls, the
direction and the degree of misclassification were similar for cases and controls. Thus, we found little evidence of bias related to case status. Our conclusions were not altered if only cases and controls with a complete questionnaire response were analysed. A regression analysis indicated that the amount of difference between questionnaire and interview response did depend on case-control status for some foods, however. During interview, cases tended to overestimate low levels of consumption and to underestimate high levels of consumption reported on the questionnaire for meat, snacks, and coffee and tea. Notably, this difference in reporting did not occur for the main food groups discussed in the media as being related to breast cancer.

The weak agreement between the questionnaire and interview responses was not unexpected given several differences in method of reporting: the questionnaire was self-administered, it concentrated solely on current diet, and it was structured by food groups; whereas the interview was conducted over the telephone, it proceeded according to dietary periods from past diet to the present, it was meal-specific, and it contained two additional foods. The weak agreement is in itself a methodological problem in dietary epidemiology, but in this particular setting we studied if the pattern of agreement was different for cases and controls.

For all food groups, the linear regression of the monthly consumption reported during the interview on the monthly consumption reported on the questionnaire yields a line with a slope less than 1. Participants with high questionnaire food frequencies have lower interview frequencies, and vice versa. This regression result may simply be a statistical artifact from measurement error. However, in all food groups except alcoholic beverages, the standard deviation of monthly consumption from the interview is smaller than that from the questionnaire. This finding could indicate a truly systematic effect since the smaller standard deviation of the interview variable could result from smaller errors in the interview than in the questionnaire. This assumption is supported by another comparison of questionnaire and interview responses. The results for alcoholic beverages could also be accommodated. If the interview leads to more correct observations, but participants who report high alcohol intake on the questionnaire tend to report a too low value at interview, a negative correlation between actual value and error would be implied.

According to the regression analysis, we found that the food groups meat, snacks, and coffee and tea showed signs of possible recall bias. While we did not find any signs of recall bias for those main food groups that have attracted mass media attention as possible breast cancer risk factors when these data were collected, some women could possibly have identified red meat and snacks as possible sources of fat. However, if an awareness of the hypothesis under study was the source of the differential misclassification in the three implicated food groups, it is inconsistent that we did not see any signs of a recall bias concerning vegetables, alcohol intake and for fatty foods. A possible common denominator for all three food groups is that treatment factors such as convalescence after surgery, and adjuvant radiotherapy and chemotherapy might lower the appetite and thus affect recall for foods in these groups. If our regression analyses reflect such an effect, this finding can have two implications for case-control studies of diet and disease risk. We may have to learn more about how the disease under study and its treatment affects dietary patterns. Furthermore, we may have to avoid using hospital controls since they may have undergone treatments that significantly affect dietary recall. In fact this problem might be larger after e.g. bowel and urinary tract surgery where dietary counselling and altered physiology might influence dietary habits and recall more than after breast cancer treatment.

Only for vegetables did the odds ratio based on the interview differ from that based on the questionnaire. This outcome might have led to different interpretations of dietary risks had we had only one source of information available. The appearance of a protective effect of vegetables in the interview, as compared with the null effect indicated in the questionnaire, may be due to recall bias because the study subjects may have changed their recall according to a perceived hypothesis. However, this finding must be interpreted cautiously. The confidence limits for those risk estimates did overlap. The regression analysis did not indicate any recall bias for vegetables. The vegetable food group is the only food group to show this possible difference. The hypothesis of a protective effect of vegetables did not receive greater publicity in the Swedish media than discussions about fat, alcohol, and breast cancer risk during this time. Theoretically, some of the differences between cases and controls concerning the fruit and vegetable food group could be due to seasonal variations. However, in all analyses the study design with the frequency matching on month on entry into a cohort was accommodated for. Additionally, the mean time between completed questionnaire and interview was the same for cases and controls and thus effects of seasonal variations should cancel out between cases and controls.

This study did not substantiate reports that recall ability depended on type of diet and level of education.
However, we had a low power to detect differences by type of diet.

In summary, dietary recall did depend on case-control status for some food groups in our study. However, these food groups were not the ones being discussed at the time as risk factors for breast cancer either in the scientific community or in the media. These results do agree with previous investigators’ findings of weak or no signs of recall bias in breast cancer case-control studies.

ACKNOWLEDGEMENT
The analyses were undertaken while L Holmberg was a visiting scientist with the Centers for Disease Control. The study was supported by grants from the Swedish Cancer Society.

REFERENCES
12 Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968; 70: 213-20.

(Revised version received August 1995)
**APPENDIX**

*Dairy Products*
- Butter for sandwiches
- Margarine for sandwiches
- Margarine-butter mix
- Margarine, 40% fat
- Cheese
- Milk, 0.5% fat
- Milk, 1.5% fat
- Milk, 3% fat
- Cultured milk, 0.5% fat
- Cultured milk, 3% fat

*Vegetables*
- Boiled potatoes
- Fried potatoes
- French fries
- Root vegetables
- Cabbage
- Tomatoes
- Lettuce, cucumbers
- Spinach
- Vegetables, others than above

*Fruits*
- Apples, pears
- Oranges, citrus fruits
- Bananas
- Juice

*Meat*
- Meat
- Meat stew
- Bacon
- Minced meat dishes

*Sausage*  
- Meat & sausage for sandwiches
- Liver paste
- Blood pudding
- Liver, kidney
- Chicken
- Eggs, egg dishes
- Fish: salmon, mackerel, tuna
- Other fish
- Shellfish
- Soups

*Cereals*
- Coarse bread
- Bread
- Crisp bread
- Porridge, gruel
- Cereals, müsli
- Pancake
- Rice
- Pasta
- Stewed brown beans, pea soup

*Snacks*
- Potato crisps, popcorn, etc.
- Buns, cookies
- Ice cream
- Sweet soups
- Marmalade, jam
- Fruit syrup
- Soft drinks
- Sweets
- Chocolate
- Sugar

*Coffee, Tea*
- Coffee
- Tea

*Alcoholic Beverages*
- Light beer, 1.8% w/w alcohol
- Beer, 2.8% w/w alcohol
- Beer, 4.5% w/w alcohol
- Wine
- Spirits, liquor

*Fatty Foods*
- Butter for sandwiches
- Cheese
- Milk, 3% fat
- Cultured milk, 3% fat
- Fried potatoes
- French fries
- Bacon
- Sausage
- Liver paste
- Potato crisps, popcorn etc.
- Buns, cookies
- Sweets
- Chocolate
- Sugar