Dietary Fat and Coronary Heart Disease: A Comparison of Approaches for Adjusting for Total Energy Intake and Modeling Repeated Dietary Measurements


Previous cohort studies of fat intake and risk of coronary heart disease (CHD) have been inconsistent, probably due in part to methodological differences and various limitations, including inadequate dietary assessment and incomplete adjustment for total energy intake. The authors analyzed repeated assessment of diet from the Nurses' Health Study to examine the associations between intakes of four major types of fat (saturated, monounsaturated, polyunsaturated, and trans fats) and risk of CHD during 14 years of follow-up (1980-1994) by using alternative methods for energy adjustment. In particular, the authors compared four risk models for energy adjustment: the standard multivariate model, the energy-partition model, the nutrient residual model, and the multivariate nutrient density model. Within each model, the authors compared four different approaches for analyzing repeated dietary measurements: baseline diet only, the most recent diet, and two different algorithms for calculating cumulative average diets. The substantive results were consistent across all models; that is, higher intakes of saturated and trans fats were associated with increased risk of CHD, while higher intakes of monounsaturated and polyunsaturated fats were associated with reduced risk. When nutrients were considered as continuous variables, the four energy-adjustment methods yielded similar associations. However, the interpretation of the relative risks differed across models. In addition, within each model, the methods using the cumulative averages in general yielded stronger associations than did those using either only baseline diet or the most recent diet. When the nutrients were categorized according to quintiles, the residual and the nutrient density models, which gave similar results, yielded statistically more significant tests for linear trend than did the standard and the partition models. Am J Epidemiol 1999; 149:531-40.

Numerous metabolic studies have demonstrated strong cholesterol-lowering effects for polyunsaturated fat when it is substituted for dietary saturated fat or carbohydrates (1). However, results from previous prospective epidemiologic studies on dietary fat intake and risk of coronary heart disease (CHD) have been inconsistent; only one study found a significant inverse association for polyunsaturated fat intake (2), and all others yielded equivocal results (3-12). This discrepancy may result from methodological limitations, including small study size and inadequate dietary assessment methods. Some studies used 24-hour recall to represent long-term exposure (4, 5, 7, 9, 10, 12), and others used dietary assessment instruments of unknown validity and reproducibility (6, 8). In addition, prior studies have not collected dietary information repeatedly during follow-up, which could be used to reduce measurement error due to intraindividual variation. Many studies did not appropriately account for total energy intake, which is important for controlling for confounding and removing extraneous variation (13). Most studies also failed to control for intake of trans fat. This may seriously confound the associations for unsaturated fats because of strong correlations between trans fat and monoo- and polyunsaturated fats. Finally, because different types of fat are intercorrelated due to the same food sources, it is important to adjust simultaneously for each other in the analyses. However, most previous analyses considered each type of fat separately.
We have reported previously on intakes of total fat and specific sources and types of fat in relation to risk of CHD among Nurses' Health Study participants (14). Briefly, we found that the intakes of saturated and trans fats were associated with increased risk, whereas the intakes of mono- and polyunsaturated fats were associated with lower risk. Owing to opposing associations of specific type of fats, total fat consumption was not associated with the risk. In this study, we focus on several methodological issues. In particular, we compare four risk models for energy adjustment (13): the standard multivariate model, the nutrient residual (energy-adjusted) model, the energy-partition model, and the multivariate nutrient density model. Within each model, we compare four different approaches for analyzing repeated dietary measurements: baseline diet only, the most recent diet, and two cumulative average diets. We also examine the effect of categorization of the nutrients on the estimates from the risk models.

**MATERIALS AND METHODS**

**Study population and endpoints**

The details of the study population have been described elsewhere (14). Briefly, the study population was 80,082 Nurses' Health Study participants aged 34-59 years and without previously diagnosed CHD, stroke, cancer, hypercholesterolemia, or diabetes in 1980. Follow-up continued for 14 years, until diagnosis of CHD, death, or June 1, 1994. Every 2 years, follow-up questionnaires have been sent to update information on potential risk factors and to identify newly diagnosed cases of CHD and other diseases. In 1980, a food frequency questionnaire was included to assess intake of fat and other nutrients. The questionnaire contained 61 foods that allowed maximum discrimination among intake of total, saturated, monounsaturated, polyunsaturated, and trans fats, and other nutrients. In 1984, the food frequency questionnaire was expanded to include 116 items. The primary change in the revised questionnaire was to create individual questions from groups of nutritionally similar foods that had previously been collapsed into single items. In a subsample of women from the main study, both the 61- and 116-item versions of the questionnaires have been shown to provide a reasonable measure of dietary intakes of saturated, monounsaturated, and polyunsaturated fats and of other nutrients when compared with four 1-week diet records (15, 16) (table 1). We sent similar expanded questionnaires to update diet in 1986 and 1990.

The outcome included nonfatal myocardial infarction and fatal CHD. Nonfatal myocardial infarction was confirmed if the criteria of the World Health Organization were satisfied, that is, symptoms plus either diagnostic electrocardiographic changes or elevated cardiac enzyme level (17). Fatal CHD was defined as fatal myocardial infarction if confirmed by hospital records or autopsy or if CHD was listed as the cause of death on the death certificate and this was the underlying and only plausible cause and evidence of previous CHD was available. During 14 years of follow-up (1980-1994), we identified a total of 658 new cases of nonfatal myocardial infarction and 281 fatal cases of CHD.

**Risk models for adjusting for total energy intake**

Our primary exposures of interest were intakes of four major types of fat: saturated, monounsaturated, polyunsaturated, and trans fats, which were all correlated with total energy intake (correlation coefficients ranged from 0.72 for trans fat to 0.84 for saturated fat). We used four different methods to adjust for total energy (13):

1) Standard multivariate model, in which the independent exposures are absolute intake of fats (grams per day) and total energy intakes. The relation can be expressed as the following:

\[
\text{Disease risk} = s_1 \times \text{saturated fat} + s_2 \times \text{mono} + s_3 \times \text{poly} + s_4 \times \text{trans} + s_5 \times \text{protein} + s_6 \times \text{total energy intake},
\]

where \(s_1-s_6\) are regression coefficients.

We included protein intake in the model so that the coefficients for one type of fat can be interpreted as the effect of substituting a certain amount of this fat (e.g., 10 g) for the same amount of energy from carbohydrates while holding constant the intakes of total energy, protein, and other fats. It should be noted that the coefficient for total energy in this model does not represent the effect of the energy, but rather the effect of carbohydrates.

**TABLE 1. Pearson correlation coefficients for energy-adjusted nutrients between 1980 and 1984 dietary questionnaires and 1980 diet records**

<table>
<thead>
<tr>
<th></th>
<th>1980 compressed questionnaire vs. 1980 diet records</th>
<th>1980 compressed questionnaire vs. 1980 diet records</th>
<th>1980 revised questionnaire vs. 1980 diet records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total fat</td>
<td>0.45</td>
<td>0.52</td>
<td>0.54</td>
</tr>
<tr>
<td>Saturated fat</td>
<td>0.47</td>
<td>0.57</td>
<td>0.53</td>
</tr>
<tr>
<td>Monounsaturated fat</td>
<td>0.51</td>
<td>0.56</td>
<td>0.48</td>
</tr>
<tr>
<td>Polyunsaturated fat</td>
<td>0.51</td>
<td>0.46</td>
<td>0.58</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0.52</td>
<td>0.58</td>
<td>0.57</td>
</tr>
<tr>
<td>Protein</td>
<td>0.53</td>
<td>0.44</td>
<td>0.52</td>
</tr>
<tr>
<td>Total carbohydrate</td>
<td>0.44</td>
<td>0.39</td>
<td>0.61</td>
</tr>
</tbody>
</table>
2) Nutrient residual model, in which the fat residuals obtained by regressing nutrient intake on total energy intake (18) are included as independent variables. Total energy intake is also included as a covariate. The relation can be expressed as the following:

\[
\text{Disease risk} = n1 \cdot \text{saturated fat residual} + n2 \cdot \text{monounsaturated fat residual} + n3 \cdot \text{polyunsaturated fat residual} + n4 \cdot \text{trans fat residual} + n5 \cdot \text{protein residual} + n6 \cdot \text{total energy intake}.
\]

The coefficients for the nutrient residuals have the same isocaloric substitution interpretation as does the standard multivariate model, and they are quantitative-ly the same (19). The coefficient for energy, however, retains the biologic meaning of total energy intake.

Because energy intake and the nutrient intake are typically skewed toward higher values and the variation in nutrient intake (and, thus, the residuals) is greater at higher total energy intake, we used log transformation of the dietary intake variables to create residuals with a more constant variance across the levels of total energy intake (18). Since nutrients expressed as residuals have an unfamiliar scale (and even more so when logarithmically transformed), a back-transformation can be made by adding a constant (e.g., the predicted value for log of 2,000 calories) and then taking the antilog. Because of the use of log transformation in the calculation of the residuals, the quantitative equivalence of the coefficients for dietary fats from the standard multivariate and nutrient residual models may not hold.

3) Energy-partition model, in which total energy is partitioned into that contributed by different types of fats and that contributed by other sources (protein and carbohydrates). Because units as either calories or grams for the nutrients do not in any way affect the analyses, we model absolute intakes of all energy-bearing nutrients simultaneously:

\[
\text{Disease risk} = e1 \cdot \text{saturated fat} + e2 \cdot \text{monounsaturated fat} + e3 \cdot \text{polyunsaturated fat} + e4 \cdot \text{trans fat} + e5 \cdot \text{protein} + e6 \cdot \text{carbohydrates}.
\]

In this model, the coefficient for one type of fat represents the effect of increasing absolute intake of this type of fat while holding constant other fats and other energy sources (i.e., protein and carbohydrates). Therefore, it represents the effect of "adding fat," which includes both its energy and nonenergy effects.

4) Multivariate nutrient density model, in which the nutrient densities (percentages of energy) from different types of fat are included as exposure variables and total energy is included as a covariate:

\[
\text{Disease risk} = m1 \cdot \text{percent energy from saturated fat} + m2 \cdot \text{percent energy from monounsaturated fat} + m3 \cdot \text{percent energy from polyunsaturated fat} + m4 \cdot \text{percent energy from trans fat} + m5 \cdot \text{percent energy from protein} + m6 \cdot \text{total energy}.
\]

The coefficients from this model also have an isocaloric interpretation, representing substitution for an equal energy from carbohydrate, but are in units of the percentage of energy.

**Methods for analyzing repeated dietary measurements**

The statistical model for the analyses was the pooled logistic regression model originally proposed by Wu and Ware (20) and further discussed by Cupples et al. (21) and D'Agostino et al. (22). In this approach, each 2-year examination interval is treated as an independent follow-up study. Observations over all intervals are pooled into a single sample, and a logistic regression is used to examine the association between the risk factors and development of disease during all follow-up periods. D'Agostino et al. (22) showed the asymptotic equivalence of this approach and the Cox regression model with time-dependent covariates. The necessary conditions for this equivalence include relatively short time intervals and small probability of the outcome in the intervals, both of which are satisfied here. Following D'Agostino et al., a general pooled logistic model can be written:

\[
\text{Logit} p_i(X(t_{i-1})) = \alpha_i + \beta_1 x_1(t_{i-1}) + \ldots + \beta_p x_p(t_{i-1})
\]

where \( p_i(X(t_{i-1})) \) is the conditional probability of observing an event by time \( t_i \) given that the individual is event-free at time \( t_{i-1} \), and \( X(t_{i-1}) \) is the vector of independent variables measured at time \( t_{i-1} \).

The above model assumes that only the current risk profile is needed to predict disease outcome. For dietary exposures, previous diet or long-term diet may be more important. Therefore, we compared the following approaches for handling repeated dietary measurements in the pooled logistic regression:

1) Baseline diet only, in which CHD incidence from 1980 to 1994 was related to 1980 fat intake.
2) Most recently measured diet, in which CHD incidence during the 1980-1984 time period was related to fat intake from the 1980 questionnaire, CHD incidence during the 1984-1986 time period
was related to fat intake from the 1984 questionnaire, CHD incidence during the 1986–1990 time period was related to fat intake from the 1986 questionnaire, and CHD incidence during the 1990–1994 time period was related to fat intake from the 1990 questionnaire.

3) Cumulative average diet, in which CHD incidence between each 2-year questionnaire cycle was related to the cumulative average of fat intake calculated from all of the preceding dietary measures. Therefore, CHD incidence during the 1980–1984 time period was related to the fat intake from the 1980 questionnaire; CHD incidence during the 1984–1986 time period was related to the average intake from the 1980 and 1984 questionnaires; CHD incidence during the 1986–1990 time period was related to the average intake from the 1980, 1984, and 1986 questionnaires; and CHD incidence during the 1990–1994 time period was related to the average intake from all four dietary questionnaires.

4) As an alternative, to give more weight to the most recent diet, we calculated the average of the most recently measured diet and previous average diet. For 1980–1984 and 1984–1986 time periods, the exposures were calculated in the same way as described above. For CHD incidence during 1986–1990 time period, the exposures were calculated as follows:


Similarly, for CHD incidence that occurred during the 1990–1994 time period, the exposures were calculated as follows:


In both cumulative methods, we used the averages of fat intakes instead of a single measurement to reduce within-subject variation and best represent long-term diet. One method provides more weight to the baseline diet, while the other assumes that recent diet is more important.

**Change of diet due to intermediate events**

Because change in diet after development of intermediate endpoints, such as hypercholesterolemia, hypertension, diabetes, or angina, may confound the exposure-disease association (23), we stopped updating diet at the time interval during which individuals developed those intermediate endpoints in the updated analyses (using the most recent diet or the cumulative averages).

**Missing data**

On average, about 80 percent of baseline population completed the repeated dietary questionnaires during each separate follow-up cycle. Nonresponse is not an issue in the analyses using baseline diet only because only women with complete baseline dietary data were included in all analyses. In the analyses using the most recent diet, we used last observation carried forward method (24) to impute the missing values. For example, we used nutrient values from the 1980 questionnaire to replace missing values in 1984 and then used 1984 values to replace 1986 missing values, and so on. Missing data indicators were assigned to those who had missing values in 1984, 1986, or 1990. In the analyses using the cumulative averages, we calculated the average intake from all available dietary questionnaires. Again, missing data indicators were assigned for each time period to those who had missing data. Alternative analyses eliminating subjects with missing data yielded similar associations, but with wider confidence intervals.

**Nondietary covariates**

Nondietary covariates, including age, cigarette smoking, body mass index, menopausal hormone use, alcohol use, multivitamin use, and vitamin E supplement use were updated biennially. Aspirin use was assessed in 1980, 1982, 1984, and 1988. Vigorous exercise was assessed in 1980. The analyses using baseline or updated nondietary covariates yielded similar results. Here we report analyses using updated covariates.

**RESULTS**

At baseline, the mean intakes were 15.6 percent of energy from saturated fat, 16.0 percent from monounsaturated fat, 5.3 percent from polyunsaturated fat, and 2.2 percent from trans fats. Figure 1 shows time trends of dietary fat intakes. Figure 2 shows changes in fat intake from 1980 to 1984 according to whether a person developed an intermediate event. The reduction in intakes of saturated, monounsaturated, and trans fats was greatest for those who developed angina; intermediate for those who developed diabetes, hypertension, or hypercholesterolemia; and least for those who did not have those events, whereas the increase in polyunsaturated fat intake was smallest for those who developed angina. This suggests that development of those intermediate endpoints resulted in change in diet. Therefore, in the
analyses using updated dietary information, we stopped updating diet at the beginning of the time interval when an individual developed an intermediate endpoint.

Table 2 shows correlation coefficients for each type of fat and total fat between any two dietary assessments. The correlation coefficients ranged from 0.23.
to 0.59, depending on the closeness in the time of measurements.

Table 3 shows relative risks of CHD and 95 percent confidence intervals estimated from various risk models using baseline diet, most recent diet, and two cumulative average diets. We estimated two sets of

| TABLE 3. RR* and 95% CI† of coronary heart disease for intakes of major types of fat (g/day) estimated from various models for energy adjustment and handling repeated measurements‡ |
|-----------------|-----------|-----------|-----------|-----------|-----------|
|                  | Baseline | Most recent | Cumulative | Cumulative |
|                  | diet only|          | average [§] | average [¶] |           |
|                  | RR  95% CI | RR  95% CI | RR  95% CI | RR  95% CI |           |
| Standard multivariate model (without calories in the model) | | | | | |
| Saturated fat (10 g) | 1.17 1.00–1.36 | 1.19 1.02–1.40 | 1.18 0.99–1.40 | 1.18 0.99–1.40 |
| Monounsaturated fat (10 g) | 0.85 0.71–1.03 | 0.88 0.72–1.07 | 0.84 0.67–1.04 | 0.84 0.68–1.05 |
| Polyunsaturated fat (10 g) | 0.74 0.55–1.01 | 0.87 0.70–1.07 | 0.62 0.46–0.83 | 0.66 0.49–0.87 |
| Trans fat (5 g) | 1.65 1.28–2.68 | 1.49 1.06–2.07 | 2.32 1.56–3.45 | 2.24 1.52–3.31 |
| Total fat (10 g) | 1.03 | 0.99–1.07 | 1.04 | 1.01–1.08 | 1.03 | 0.99–1.09 | 1.04 | 0.99–1.09 |
| Standard multivariate model (with calories in the model) | | | | | |
| Saturated fat (10 g) | 1.15 0.97–1.37 | 1.17 0.98–1.39 | 1.17 0.96–1.42 | 1.17 0.97–1.42 |
| Monounsaturated fat (10 g) | 0.86 0.71–1.04 | 0.88 0.72–1.08 | 0.84 0.67–1.06 | 0.85 0.68–1.05 |
| Polyunsaturated fat (10 g) | 0.73 0.53–1.00 | 0.85 0.67–1.08 | 0.61 0.45–0.84 | 0.65 0.48–0.88 |
| Trans fat (5 g) | 1.83 1.26–2.66 | 1.48 1.07–2.06 | 2.31 1.54–3.45 | 2.23 1.50–3.31 |
| Total energy (100 kcal) | 1.01 | 0.97–1.04 | 1.01 | 0.98–1.04 | 1.00 | 0.97–1.04 | 1.00 | 0.97–1.04 |
| Total fat (10 g) | 1.01 | 0.97–1.06 | 1.03 | 0.98–1.08 | 1.03 | 0.98–1.08 |
| Nutrient residual model | | | | | |
| Saturated fat (10 g) | 1.16 0.96–1.39 | 1.16 0.94–1.42 | 1.18 0.96–1.45 | 1.19 0.96–1.46 |
| Monounsaturated fat (10 g) | 0.83 0.68–1.02 | 0.78 0.62–0.99 | 0.80 0.63–1.02 | 0.80 0.63–1.02 |
| Polyunsaturated fat (10 g) | 0.71 0.51–0.99 | 0.76 0.58–1.00 | 0.58 0.41–0.82 | 0.61 0.44–0.86 |
| Trans fat (5 g) | 2.00 1.37–2.92 | 1.93 1.30–2.87 | 2.51 1.66–3.80 | 2.45 1.53–3.68 |
| Total energy (100 kcal) | 1.01 1.00–1.03 | 1.01 | 0.99–1.03 | 1.01 | 0.99–1.03 |
| Total fat (10 g) | 1.01 0.96–1.06 | 1.01 | 0.96–1.07 | 1.00 | 0.94–1.06 |
| Energy-partition model** | | | | | |
| Saturated fat (10 g) | 1.15 0.98–1.35 | 1.17 0.99–1.38 | 1.16 0.97–1.40 | 1.17 0.97–1.40 |
| Monounsaturated fat (10 g) | 0.87 0.71–1.07 | 0.89 0.73–1.00 | 0.86 0.68–1.08 | 0.86 0.68–1.09 |
| Polyunsaturated fat (10 g) | 0.73 0.53–0.99 | 0.85 0.68–1.08 | 0.81 0.44–0.83 | 0.64 0.47–0.87 |
| Trans fat (5 g) | 1.82 | 1.25–2.65 | 1.49 | 1.07–2.07 | 2.24 | 1.53–3.42 | 2.21 | 1.49–3.28 |
| Total fat (10 g) | 1.03 0.99–1.06 | 1.04 0.99–1.08 | 1.03 0.99–1.07 | 1.04 1.00–1.08 |
| Multivariate nutrient density model | | | | | |
| Saturated fat (5% energy) | 1.14 0.96–1.34 | 1.19 | 1.00–1.42 | 1.17 0.97–1.41 | 1.18 0.98–1.42 |
| Monounsaturated fat (5% energy) | 0.85 0.71–1.02 | 0.80 | 0.66–0.98 | 0.82 | 0.68–1.02 | 0.83 | 0.67–1.03 |
| Polyunsaturated fat (5% energy) | 0.74 0.55–1.00 | 0.84 | 0.67–1.06 | 0.83 | 0.47–0.86 | 0.67 | 0.50–0.90 |
| Trans fat (2% energy) | 1.63 1.24–2.14 | 1.58 | 1.21–2.07 | 1.93 | 1.43–2.60 | 1.85 | 1.38–2.48 |
| Total energy (100 kcal) | 1.01 1.00–1.02 | 1.01 | 1.00–1.02 | 1.01 | 0.99–1.02 | 1.01 | 0.99–1.02 |
| Total fat (5% energy) | 1.01 0.97–1.05 | 1.01 | 0.97–1.06 | 1.02 | 0.97–1.07 | 1.02 | 0.97–1.07 |

* Relative risks (RRs) adjusted for age (5-year category); time period (seven periods); body mass index (five categories); cigarette smoking (never, past, and current smoking of 1–14, 15–24, and ≥25 cigarettes per day); menopausal status (premenopausal, postmenopausal without hormone replacement, postmenopausal with past hormone replacement, and postmenopausal with current hormone replacement); parental history of myocardial infarction before age 60 years; multiple vitamin use; vitamin E supplement use; alcohol consumption (four categories); history of hypertension; aspirin use (nonuser, 1–6 weeks, ≥7 weeks, and dosage unknown); vigorous exercise ≥1 time/week; cholesterol intake; and protein intake. Energy-partition model included carbohydrate intake, and other models included total energy intake. Except for models using baseline diet only, all models included three missing data indicators for those who did not complete follow-up dietary questionnaires.

† CI, confidence interval.
‡ All dietary variables including energy intake were entered into the models in continuous forms.
§ See method 3 for analyzing repeated dietary measurements.
¶ See method 4 for analyzing repeated dietary measurements.
# Total fat was entered into a separate model, which did not include specific types of fats.
** Total energy intake was not included in the model.
standard multivariate models, one without the energy term and the other including total energy. We also ran separate models for types of fat (saturated, monounsaturated, polyunsaturated, and trans fats in one model) and total fat. In all of these models, dietary variables, including total energy intake, were treated as continuous variables. In the standard multivariate models including the energy term, total energy was not significantly associated with CHD risk. As discussed previously, the association for total energy in these models should be interpreted as the effect of carbohydrates. In the residual model, in which the effect of total energy can be interpreted, the association for total energy was not significant either. The results for types of fat and total fat were similar across all models. Specifically, intakes of saturated and trans fats were associated with increased risk. For polyunsaturated and saturated fats were associated with reduced risk. Total fat was generally not significantly associated with the risk. For specific types of fats, within each risk model, the models using the cumulative averages yielded associations similar to each other, and the models using either baseline diet only or most recent diet gave results similar to each other. In general, the models using cumulative averages yielded stronger associations, especially for polyunsaturated and trans fats.

Next we categorized the continuous nutrient intakes in the above models into quintiles and reran all of the models. In all models, total energy was also categorized into quintiles. Since the general pattern of the results was similar for different approaches to using the repeated dietary measurements, we reported only the results from the analyses using the second cumulative average diet (table 4). Again, the associations for the four major types of fat were qualitatively similar across the four models. However, the magnitude of the associations was more similar between the standard and the energy-partition models and between the residual and the nutrient density models. For polyunsaturated fat, the trend of decreasing risk with higher intake was significant in the residual and the nutrient density models, but not in other two models. For trans fat, the trend of increasing risk with higher intake was highly significant in the residual and the nutrient density models, but only marginally significant in other two models. Interestingly, the relative risk for the highest quintile of trans fat intake was almost identical across the four models, but the 95 percent confidence interval was wider in the standard model (relative risk (RR) = 1.47, 95 percent confidence interval (CI) 1.02–2.12) and the partition model (RR = 1.46, 95 percent CI 1.01–2.11), compared with the residual model (RR = 1.46, 95 percent CI 1.10–1.93) and the nutrient density model (RR = 1.46, 95 percent CI 1.10–1.93).

### TABLE 4. RR* and 95% CI† of coronary heart disease according to quintiles of the cumulative averages of fat intake‡ from four risk models for energy adjustment

<table>
<thead>
<tr>
<th>Measurement</th>
<th>1 (lowest)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 (highest)</th>
<th>p for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR</td>
<td>95% CI</td>
<td>RR</td>
<td>95% CI</td>
<td>RR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Standard multivariate model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated fat</td>
<td>1.00</td>
<td>0.82–1.44</td>
<td>1.38</td>
<td>0.97–1.95</td>
<td>1.28</td>
<td>0.84–1.95</td>
</tr>
<tr>
<td>Monounsaturated fat</td>
<td>1.00</td>
<td>0.83–1.14</td>
<td>0.80</td>
<td>0.54–1.17</td>
<td>0.82</td>
<td>0.51–1.31</td>
</tr>
<tr>
<td>Polyunsaturated fat</td>
<td>1.00</td>
<td>0.80–1.18</td>
<td>0.93</td>
<td>0.75–1.16</td>
<td>0.96</td>
<td>0.77–1.20</td>
</tr>
<tr>
<td>Trans fat</td>
<td>1.21</td>
<td>0.95–1.55</td>
<td>1.25</td>
<td>0.95–1.66</td>
<td>1.39</td>
<td>1.02–1.90</td>
</tr>
<tr>
<td>Nutrient residual model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated fat</td>
<td>1.00</td>
<td>0.81–1.33</td>
<td>0.97</td>
<td>0.74–1.28</td>
<td>1.22</td>
<td>0.91–1.64</td>
</tr>
<tr>
<td>Monounsaturated fat</td>
<td>1.00</td>
<td>0.75–1.24</td>
<td>1.08</td>
<td>0.81–1.44</td>
<td>0.95</td>
<td>0.69–1.31</td>
</tr>
<tr>
<td>Polyunsaturated fat</td>
<td>1.00</td>
<td>0.75–1.12</td>
<td>0.82</td>
<td>0.66–1.02</td>
<td>0.85</td>
<td>0.68–1.08</td>
</tr>
<tr>
<td>Trans fat</td>
<td>1.03</td>
<td>0.82–1.29</td>
<td>1.10</td>
<td>0.86–1.40</td>
<td>1.28</td>
<td>1.00–1.65</td>
</tr>
<tr>
<td>Energy-partition model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated fat</td>
<td>1.00</td>
<td>0.80–1.41</td>
<td>1.33</td>
<td>0.94–1.88</td>
<td>1.21</td>
<td>0.80–1.84</td>
</tr>
<tr>
<td>Monounsaturated fat</td>
<td>1.00</td>
<td>0.62–1.12</td>
<td>0.78</td>
<td>0.53–1.15</td>
<td>0.80</td>
<td>0.50–1.27</td>
</tr>
<tr>
<td>Polyunsaturated fat</td>
<td>1.00</td>
<td>0.79–1.16</td>
<td>0.92</td>
<td>0.74–1.14</td>
<td>0.85</td>
<td>0.76–1.19</td>
</tr>
<tr>
<td>Trans fat</td>
<td>1.21</td>
<td>0.85–1.54</td>
<td>1.24</td>
<td>0.94–1.84</td>
<td>1.38</td>
<td>1.01–1.88</td>
</tr>
<tr>
<td>Multivariate nutrient density model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated fat</td>
<td>1.00</td>
<td>0.73–1.19</td>
<td>0.68</td>
<td>0.65–1.33</td>
<td>1.09</td>
<td>0.81–1.45</td>
</tr>
<tr>
<td>Monounsaturated fat</td>
<td>1.00</td>
<td>0.83–1.38</td>
<td>1.10</td>
<td>0.82–1.48</td>
<td>0.98</td>
<td>0.71–1.37</td>
</tr>
<tr>
<td>Polyunsaturated fat</td>
<td>1.00</td>
<td>0.76–1.14</td>
<td>0.86</td>
<td>0.69–1.07</td>
<td>0.86</td>
<td>0.68–1.08</td>
</tr>
<tr>
<td>Trans fat</td>
<td>1.02</td>
<td>0.81–1.28</td>
<td>1.12</td>
<td>0.88–1.42</td>
<td>1.24</td>
<td>0.97–1.80</td>
</tr>
</tbody>
</table>

* Relative risks (RRs) adjusted for the same covariates as in table 3.
† CI, confidence interval.
‡ See method 4 for analyzing repeated dietary measurements.
DISCUSSION

We compared four alternative approaches for adjustment for total energy intake: standard, residual, partition, and nutrient density. Within each approach, we compared four different methods for analyzing repeated dietary measurements: baseline diet only, the most recent diet, and two cumulative average diets. When the nutrients were in their continuous form, the four energy-adjustment methods yielded similar associations. When the nutrients were categorized according to quintiles, the residual and the nutrient density models, which gave similar results, yielded narrower confidence intervals for the relative risks and statistically more significant tests for linear trend than did the standard and the partition models. In general, the models using the cumulative averages yielded stronger associations than did those using either only baseline diet or most recent diet. Despite the methodological differences, all of the models yielded similar substantive results; that is, saturated and trans fats are associated with increased risk of CHD, whereas mono- and polyunsaturated fats are associated with lower risk.

In the past decade, there has been considerable debate over the most appropriate method for accounting for total energy intake in the analysis (25). Four methods have been proposed and used in various analyses: standard, residual, partition, and nutrient density. Although the first three methods are mathematically equivalent when the nutrients are in their continuous form (19, 26–28), the interpretation of the coefficients is not the same across the models. For the standard and residual models, the estimated effects for specific types of fat have "isocaloric substitution" interpretation. For example, the relative risk of 2.23 for trans fat in the standard model with the second cumulative average can be interpreted as a 123 percent increase in risk associated with substitution of 5 g of trans fat for the same amount of energy from carbohydrates. Note that the relative risks estimated from the corresponding standard multivariate and residual models are similar but not identical, as one would expect due to the use of logarithm transformed nutrients and calories in the regression analyses for energy adjustment. Despite the same interpretation of the relative risks for dietary fats, the interpretation of the associations for total energy is different between the standard and residual models. As discussed previously, the association for total energy in the standard model should be interpreted as the effect of carbohydrates, whereas it can be interpreted as the effect of total energy intake in the residual model.

The relative risk of 2.21 in the corresponding partition model can be interpreted as a 121 percent increase in risk associated with a 5-g increase in trans fat intake, while holding constant intakes of other fats, protein, and carbohydrates. In this situation, total energy is also increased, and the estimated coefficient reflects both the energy and nonenergy effects of the nutrient. This is not a problem for nutrients that contribute only a small amount of total energy (e.g., trans fat) or when total energy is not related to the disease (which is the case in this example). However, when total energy intake is associated with disease, coefficients obtained from a partition model can be misleading because nutrients (e.g., saturated fat) that contribute to total energy intake may be associated with risk of the disease simply due to their nonspecific contribution to total energy intake. Nevertheless, a "substitution" coefficient can be calculated from the estimates provided by the partition model (27). Finally, the estimated relative risk of 1.85 in the corresponding nutrient density model can be interpreted as an 85 percent increase in risk associated with substitution of 2 percent of energy from trans fat for an equivalent amount of energy from carbohydrates. This interpretation is useful because public health recommendations are generally expressed in the same units.

In many circumstances, dietary variables are categorized according to quartiles or quintiles to avoid misspecification of the model and to reduce the influences of outliers. Unfortunately, the statistical equivalence among the standard, residual, and partition models no longer exists once the nutrient has been categorized (29). This was observed by Kushi et al. (30) in a study on fat intake and risk of postmenopausal breast cancer. In our analyses, although the qualitative effect for each type of fat after categorization was still similar across the risk models, the residual and the nutrient density models provided more statistically significant trends in relative risks, particularly for polyunsaturated and trans fats, than did the standard and the residual models. In addition, for relative risks of similar magnitude, the 95 percent confidence intervals were considerably narrower for the residual and the nutrient density models compared with the other two models. This result is consistent with a simulation study by Brown et al. (29). To address potential misclassification in the middle three quintiles, we reran the standard multivariate and nutrient residual models after collapsing the middle three quintiles. The results were very similar. For example, in the standard model, the relative risk for the fifth quintile of trans fat was 1.34 (95 percent CI 0.97–1.86, p for trend = 0.14). In the residual model, the relative risk was 1.39 (95 percent CI 1.06–1.82, p for trend = 0.01).

We observed considerable changes in intakes of saturated fat, monounsaturated fat, and polyunsaturated fat from 1980 to 1984 (figure 1). This may partially be
an artifact reflecting changes in the dietary questionnaires. However, these changes are generally consistent with diet record data from two validation studies conducted in 1980 and 1986, between which intake of saturated fat decreased by 12 percent, monounsaturated fat decreased by 8 percent, and polyunsaturated fat increased by 13 percent (Sampson et al., unpublished manuscript).

Null findings in diet and CHD analyses have been attributable to intraindividual variation in diet (31). One way to reduce measurement error due to intraindividual variation is to obtain repeated measurements. However, most studies to date have collected only baseline diet (corresponding to figure 3, design I). The analyses using baseline diet assume that diets are constant over time or that the rank order of dietary variables is the same across time if changes do occur (32). It is clear from this study that diets were indeed changing over time and that the changes were not random because they were affected by development of intermediate endpoints. Several previous studies have collected multiple diets at the baseline period (corresponding to figure 3, design II). For example, the Western Electric Study (2) collected dietary information at baseline and 1 year later and used the average of these two diets to predict subsequent disease risk. However, the dietary information was not further updated during 2 decades of follow-up. In the Nurses' Health Study, dietary information was collected periodically during the follow-up (figure 3, design III). One advantage of this design is that by updating dietary information in the analyses, one can accommodate changes in dietary habits among participants and in food composition. We found that the analyses using the cumulative averages yielded stronger associations than those using either only baseline diet or most recent diet. It is likely that use of the cumulative averages reduced measurement error due to intraindividual variations over time and therefore reduced attenuation of the regression coefficients. It is also possible that the cumulative averages, which reflect long-term diet, are more relevant etiologically than either the most remote (baseline) or the most recent diets. However, our analyses cannot separate error reduction from greater etiologic relevance. In addition, there are probably other sensible ways to update diet that need to be explored in further studies.

In summary, our empirical comparisons of methods for energy adjustment suggest that although qualitatively similar conclusions would be drawn from various analyses, the strength and interpretation of the findings were not the same across different methods. By analyzing the data using alternative approaches for energy adjustment, while paying careful attention to the interpretation of the coefficients, we can gain maximal insight into diet-disease relations. In addition, our results suggest the advantage of using repeated dietary measurements.

ACKNOWLEDGMENTS

Supported by research grants HL24074, HL34594, CA40356, and DK 46200 and the nutrition training grant T32DK07703 from the National Institute of Health.

The authors are indebted to Al Wing, Mark Shneyder, Stefanie Parker, Gary Chase, Karen Corsano, Lisa Dunn, Barbara Egan, and Jill Arnold for their unfailing help and to Drs. Frank E. Speizer and Charles Hennekens for their support.

REFERENCES


