Assessment of Questionnaire Validity for Measuring Total Fat Intake using Plasma Lipid Levels as Criteria

Walter Willett,1–3 Meir Stampfer,1–3 Nain-Feng Chu,3 Donna Spiegelman,3,4 Michelle Holmes,2 and Eric Rimm1–3

The validation of dietary total fat measurements has been elusive because no specific biomarker exists. In metabolic studies with controlled diets, plasma fasting triglyceride levels are reduced with higher fat intake and can thus serve as an "alloyed gold" standard. Participants in this cross-sectional analysis were 269 men aged 47–83 years from the Health Professionals Follow-up Study who completed a semiquantitative food frequency questionnaire and provided fasting blood specimens in 1994. In a multiple regression analysis adjusted for age; smoking; alcohol consumption; physical activity; body mass index; and intakes of protein, dietary fiber, and total energy, total fat intake was inversely associated with fasting triglycerides (95% confidence interval: –3.7 to –1.3%, p = 0.0002). For reported fat intakes of 20% or less of energy, the geometric mean fasting triglyceride level was 179, and for more than 40% of energy, it was 102 mg/dl. In addition, as predicted by metabolic studies, the inverse association between dietary fat and fasting triglyceride level was much stronger among overweight men than among men with a BMI of less than 25. These data provide additional evidence that informative measurements of dietary fat can be obtained by carefully constructed food frequency questionnaires. Am J Epidemiol 2001;154:1107–12.

diet; dietary fats; dietary fiber; dietary proteins; diet surveys; energy intake; questionnaires

For the last 2 decades, a reduction in total fat intake as a percentage of energy has been at the top of most dietary recommendations (1, 2). The belief that replacing dietary fat with carbohydrates would be beneficial has been based mainly on international comparisons suggesting strong epidemiologic correlations between dietary fat and risk of breast and several other cancers (1, 3). Although most case-control studies have found no significant association between dietary fat and breast cancer, a weak, but statistically significant, association was seen in a pooled analysis of primary data from these investigations (4). However, in prospective studies, little or no association has consistently been observed (5–12), and no evidence of a positive association was found in a pooled analysis of prospective studies that included nearly 5,000 cases (13). Some epidemiologists have argued that a positive association was not observed in the prospective data because the variation in fat intake in the population was insufficient or that the errors associated with the questionnaire measurements of fat intake were too large (14–16). Several lines of evidence do not support these explanations. First, specific types of fat have predicted risk of coronary heart disease within the same cohorts as expected by the effects of these fats on blood lipids in controlled metabolic studies (17, 18). In addition, the validity of dietary fat assessment has been evaluated in substudies within these cohorts by comparison with dietary records or 24-hour diet recalls; correlations have generally ranged from 0.5 to 0.7 when adjusted for energy intake (19–24). When measurement error correction methods were used to correct relative risks and confidence intervals, the 95 percent confidence intervals readily excluded the magnitude of association between dietary fat and breast cancer predicted by the ecologic comparisons (13, 25). Some have hypothesized that correlated errors in the food frequency questionnaires and the dietary records or recalls could have overstated the validity of the questionnaires in studies comparing these methods and that this might account for an observed lack of association between total fat intake and breast cancer (26, 27).

The validity of dietary assessment methods can also be assessed by comparisons with biochemical markers of intake. This is attractive mainly because the sources of error should be independent; documentation of a clear association provides strong evidence of validity of the dietary assessment method. By comparison of calculated intakes with fatty acid levels in subcutaneous adipose tissue, the validity of dietary questionnaires has been supported in many studies for measuring intakes of specific fatty acids that are not...
endogenously synthesized (23). However, a specific biochemical indicator for total fat intake does not exist, which has limited the ability to assess objectively the validity of total fat intake measurements.

An alternative approach to assess the validity of dietary total fat measurements is to use a biochemical indicator that is sensitive to intake, although not necessarily specific for fat intake. Fasting plasma triglyceride levels are reduced by increasing the percentage of energy from fat in the diet (with a corresponding decrease in percentage of energy from carbohydrates) and, thus, can serve as such a marker. In a meta-analysis of 27 controlled metabolic studies, Mensink and Katan (28) found that all types of fat (except for trans fatty acids) reduce fasting triglyceride levels with only small differences among them. Plasma high density lipoprotein (HDL) cholesterol levels (which do not necessarily need to be fasting levels) are increased by higher dietary fat intake, although they are less responsive than fasting triglycerides (28). For example, in a large, 1-year randomized trial, Knopp et al. (29) found that reducing dietary fat from 28 to 22 percent of energy increased plasma fasting triglycerides by 30 percent and reduced HDL cholesterol by 5 percent. A documented relation between fat intake calculated by dietary questionnaire and blood levels of these lipid fractions would provide evidence for the physiologic relevance of the differences in diet within a population as well as validity of the questionnaire. We therefore examined these associations in a study of men.

**MATERIALS AND METHODS**

**Population**

The Health Professionals Follow-up Study is a prospective cohort designed to investigate the association between diet and chronic disease among men. At baseline in 1986, 51,529 male US dentists, pharmacists, veterinarians, podiatrists, and osteopathic physicians aged 40–75 years completed a detailed questionnaire assessing usual dietary intake, lifestyle characteristics, and medical history. Every 2 years, participants are recontacted to update exposure information and ascertain newly diagnosed disease. Between 1993 and 1995, in addition to biennial questionnaire information, 18,225 participants provided blood samples. There was no attempt to coordinate the dates or the order of questionnaire completion and blood collection, but these were usually within 1 year of each other. The distributions of dietary and lifestyle characteristics were similar between the two populations who either did or did not return a blood sample, although men who gave blood were, on average, 1 year younger.

The original purpose of this data set was to examine the effects of alcohol consumption on blood lipids (30). For that reason, from the 18,225 men who returned a blood sample between 1993 and 1995, we excluded 8,922 men who did not have complete questionnaire information on diet, cigarette smoking, alcohol consumption, and physical activity from 1986 to 1994. Because disease conditions may influence body weight and dietary and lifestyle patterns, we also excluded 208 men with known cardiovascular disease and diabetes, gastric or duodenal ulcer, liver disease, and cancer (except nonmelanoma skin cancer). From the remaining men, we randomly sampled 467 men aged 47–83 years on the basis of seven clusters defined by their self-reported alcohol consumption pattern (e.g., abstain, light with meals, light not with meals, heavy with meals, heavy not with meals, light or heavy sometimes with meals, binge). Of these, 269 men provided a fasting (time since last meal before blood draw ≥6 hours) blood sample, and an additional 198 provided a nonfasting sample.

**Dietary information**

Average nutrient intake was derived from the semiquantitative food frequency questionnaire administered in 1994. For each food, a commonly used unit or portion size was specified, and participants were asked how often, on average, they consumed that amount of each food over the previous year. There were nine possible responses, ranging from never or less than once per month up to six or more times per day. The questionnaire also contained the types of fat commonly used and an open-ended section for foods not listed. We computed nutrient intakes by multiplying the consumption frequency of each food by the nutrient content from the US Department of Agriculture sources (31) supplemented with other data from manufacturers and published reports. Calculations also took into account information on the brand and type of margarine and types of fat used for cooking, baking, and frying at home. The dietary questionnaire has been evaluated in detail with regard to reproducibility and validity within the Health Professional’s Follow-up Study (32). The correlation for energy-adjusted total fat intake estimated by the questionnaire and by 2 weeks of weighed diet recording was 0.67 after adjustment for week-to-week variation in diet record intake, and the regression coefficient relating energy-adjusted total fat intake assessed by diet record to intake assessed by questionnaire was 0.60.

**Measurements of biochemical variables**

To reduce extraneous between-person variation, we requested fasting blood samples. However, if the subjects were not fasting, the time elapsed since the preceding meal was recorded. Blood samples were collected in three 10-ml liquid ethylenediaminetetraacetic acid blood tubes, placed on ice packs stored in insulated containers, and returned to our laboratory via overnight courier; more than 95 percent of the samples arrived within 24 hours. Upon arrival, the blood samples were centrifuged and aliquoted for storage in liquid nitrogen (−150°C). Fewer than 15 percent of the samples were slightly hemolyzed, and very few (<3 percent) were moderately hemolyzed, lipemic (<1 percent), or not cooled upon arrival (<0.5 percent).

We used standard methods to measure plasma lipids and lipoproteins (33). Plasma total cholesterol was measured by esterase-oxidase method, triglycerides were measured using an enzymatic procedure, and HDL cholesterol was assayed by the addition of magnesium ions. Coefficients of variation
were 7.4 percent for HDL cholesterol and 6.5 percent for triglycerides.

Statistical analysis

Because plasma triglyceride levels were strongly skewed to higher values, we used a natural log-normal (ln) transformation to improve normality. To assess the ability of the dietary questionnaire to predict plasma fasting ln triglyceride and HDL cholesterol levels, we used multiple regression analysis with the blood lipid level as the dependent variable and dietary fat (as percent of total energy) as a continuous independent variable. In these models, we included as covariates age; body mass index (BMI) (weight (kg)/height (m)^2); physical activity assessed by standardized questionnaire (34) (quintiles); alcohol consumption (quintiles); cigarette smoking (never, past, and current number of cigarettes/day); and dietary intakes of protein, fiber, and total energy (quintiles). These variables are known or suspected determinants of blood lipids. Because alcohol, protein, and total energy intakes are included in these multivariate nutrient-density models, the term for fat can be considered as a substitute for carbohydrate on a calorie-for-calorie basis (with total energy held constant and all energy-contributing nutrients except carbohydrate in the model, an increase in one of these implies a reciprocal decrease in carbohydrates) (23). We also created categories of total fat intake, expressed as a percentage of energy intake, and included these as dummy variables in multivariate models with the same covariates. From these models, we calculated adjusted mean lipid levels (geometric mean values for triglycerides) for each category for a nonsmoking man aged 60 years in the middle category of the other covariates.

In an earlier, small, but carefully controlled, study among postmenopausal women, the adverse changes in fasting triglyceride and HDL cholesterol levels on a high-carbohydrate/low-fat diet were larger among those with greater insulin resistance (35). Because adiposity is a major determinant of insulin resistance, we also conducted analyses stratified by BMI (<25 and ≥25 kg/m^2).

RESULTS

In this group of middle-aged and older men, mean total fat intake was 29.3 percent of energy (standard deviation = 6.6). The median fasting triglyceride level was 127 mg/dl (25th and 75th percentiles, 88 and 195), and the mean HDL cholesterol level was 60.6 mg/dl (standard deviation = 16.7).

In the multivariate models, higher total fat intake as a percent of energy, used as a continuous variable, was inversely associated with fasting ln triglyceride levels (β = −0.0249 ln mg/dl for 1 percent of energy, standard error (SE) = 0.0065, p = 0.0002). Thus, for an increase of 1 percent of energy in fat intake, triglyceride levels were lower by 2.5 percent (95 percent confidence interval (CI): −3.7, −1.3 percent). In a similar model, total fat intake was positively, but more weakly, associated with HDL cholesterol level (β = 0.237 mg/dl for 1 percent of energy, SE = 0.124, p = 0.06).

Adjusted geometric mean triglyceride and mean HDL cholesterol levels for categories of total fat intake are shown in table 1. For fat intakes 20 percent or less and greater than 40 percent of energy, triglyceride levels were 179 and 102 mg/dl, respectively, and HDL cholesterol levels were 48.8 and 52.2 mg/dl, respectively.

In analyses stratified by BMI, the inverse relation between total fat intake and fasting triglycerides was much stronger in men with a BMI of 25 or greater (figure 1). With total fat as a continuous variable in models with ln triglycerides, for BMI of less than 25, β = −0.0133, SE = 0.0088, p = 0.13 and for a BMI of 25 or more, β = −0.0422, SE = 0.0085, p < 0.0001. However, the formal test for interaction was not statistically significant. Thus, for a 1 percent increase of energy in fat intake, triglyceride levels were 1.3 percent lower (95 percent CI: −3.0, 0.4 percent) for a BMI of less than 25 and 4.1 percent lower (95 percent CI: −5.7, −2.5 percent) for BMI of 25 or greater.

DISCUSSION

In this cross-sectional analysis, higher total fat intake assessed by our self-administered food frequency questionnaire was associated with substantially lower fasting triglyceride levels; within this population, the difference in triglyceride values across levels of dietary fat as a percentage of energy was nearly twofold. A weaker and marginally significant positive association was observed between total fat intake and HDL cholesterol levels. Because these associations have been well established in metabolic studies with controlled diets (28, 29), our findings provide strong, objective evidence that physiologically important variation in dietary fat exists within this population and that the dietary questionnaire provides an informative measure of dietary total fat intake.

TABLE 1. Adjusted mean plasma fasting triglyceride and HDL cholesterol levels by percent of energy from fat†, Health Professionals Follow-up Study, 1994

<table>
<thead>
<tr>
<th>Fat intake (% energy)</th>
<th>Triglyceride‡</th>
<th>HDL cholesterol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>mg/dl</td>
</tr>
<tr>
<td>≤ 20</td>
<td>18</td>
<td>179</td>
</tr>
<tr>
<td>20.1–25</td>
<td>52</td>
<td>174</td>
</tr>
<tr>
<td>25.1–30</td>
<td>75</td>
<td>152</td>
</tr>
<tr>
<td>30.1–35</td>
<td>69</td>
<td>126</td>
</tr>
<tr>
<td>35.1–40</td>
<td>42</td>
<td>126</td>
</tr>
<tr>
<td>&gt;40</td>
<td>13</td>
<td>102</td>
</tr>
<tr>
<td>Total</td>
<td>269</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p = 0.0002§</td>
<td>p = 0.06§</td>
</tr>
</tbody>
</table>

† Adjusted for age in 1994, smoking (never, past, current <15 cigarettes/day, and current ≥15 cigarettes/day), alcohol consumption (quintiles), physical activity (quintiles), body mass index (quintiles), total energy intake (quintiles), total protein intake (quintiles), and total fat intake (quintiles). Means are adjusted to a nonsmoker aged 60 years who is in the middle category of the other covariates.
‡ Geometric means for fasting triglyceride values.
§ Test for trend using fat intake as a continuous variable.
FIGURE 1. Adjusted geometric mean plasma fasting triglyceride (TG) levels (mg/dl) by quartiles of total fat intake for men with a body mass index (BMI) of less than 25 and 25 or more kg/m², Health Professionals Follow-up Study, 1994. Data are adjusted for age in 1994, smoking, alcohol consumption, physical activity, body mass index, total energy intake, and total protein intake. The p values are from multiple regression models with total fat as a continuous variable. FFQ, food frequency questionnaire.

The strong inverse association we observed with fasting triglycerides was similar to the cross-sectional relation between dietary fat and fasting triglycerides that we previously found among postmenopausal women aged 70 years or less in the Nurses’ Health Study (36) (figure 2). In that analysis, we used the average of two dietary questionnaires, nearly identical to that used in the present population of men and administered in 1986 and 1990 to compare with blood measurements obtained in 1989. As seen in men, the association between fat intake and fasting triglycerides was strongest among those with BMIs of 25 kg/m² or more (β = −0.049 ln mg/dl for 1 percent of energy, SE = 0.012), but a significant association was still seen among women with a BMI of less than 25 (β = −0.035, SE = 0.013). The ability of the questionnaire to measure intakes of polyunsaturated fatty acids (both n-3 and n-6) and trans fatty acids has already been documented objectively by comparisons with levels in adipose tissue (37–39).

In principle, an estimate of the extent of measurement error in assessment of dietary fat from the food frequency questionnaire could be obtained by comparing the regression coefficient relating blood lipid levels to dietary fat from this study (b_{obs}) with the “true” coefficient (b_{true}) from controlled feeding studies. If no measurement error existed and the populations were similar, these coefficients should be equal. Assuming that the questionnaire measure is less than perfect, the attenuation factor for the regression of true dietary fat intake on the food frequency questionnaire value would be...
could be estimated as $\lambda = b_{obs}/b_{true}$. In the meta-analysis of controlled feeding studies in the paper by Mensink and Katan (28), polyunsaturated, monounsaturated, and saturated fat were each positively associated with serum HDL cholesterol levels when compared with the same energy intake from carbohydrates, but no data were provided for intake of total fat. We therefore estimated the true coefficient for total fat by calculating a weighted average of the coefficients for polyunsaturated, monounsaturated, and saturated fats with the weights corresponding to the proportions of these types of fat in our study population. The weighted average estimate of the true coefficient for serum HDL cholesterol, $b_{true}$, was 0.378 mg/dl for 1 percent of energy of fat substituted isocalorically for carbohydrates, which compares with our observed value of $b_{obs}$ 0.237 (95 percent CI: –0.006, 0.480). Although the observed coefficient for HDL cholesterol suggests only a modest attenuation compared to the true coefficient, i.e., $\lambda = 0.63$, the confidence interval precludes quantitative estimates of measurement error. In the same meta-analysis of controlled feeding studies with serum triglycerides (rather than ln triglycerides) as the dependent variable, Mensink and Katan (28) reported stronger, but inverse, associations for each type of fat. Because the coefficients were similar (–2.47 for polyunsaturated fat, –1.99 for monounsaturated fat, and –2.22 for saturated fat), calculation of the weighted average was not sensitive to the weights. The estimated true coefficient for serum triglycerides was –2.18 mg/dl for 1 percent of energy, which was actually smaller than the observed coefficient for untransformed serum triglyceride (–3.69, 95 percent CI: –6.47, –0.91, mg/dl for 1 percent of energy), although the observed confidence interval included the true coefficient.

One possible explanation for the higher-than-expected coefficient for serum triglycerides in our study may be differences between our participants and those in the controlled feeding trials. Insulin resistance generally increases with age as people gain fat mass and decrease physical activity, and adverse changes in blood lipids (higher triglycerides and lower HDL cholesterol) tend to be greater in the presence of greater insulin resistance (35). The men in our study were all older than age 48 years (mean age = 60.6), whereas the mean age in two thirds of feeding studies included in the meta-analyses was less than 40 years. These young adult populations tended to be much leaner than the participants in our study. These differences in age and adiposity could account for the apparently greater increases in serum triglycerides with low fat/high carbohydrate intake in our population and make a quantitative estimate of measurement error difficult. Of note, the adverse consequences of low fat intakes may have been underestimated in most feeding studies because they typically include mainly healthy young adults, whereas those at greatest risk for type 2 diabetes and coronary heart disease are primarily older adults, who tend to have greater insulin resistance.

Because insulin resistance appears to modify the relations between dietary fat and blood lipid levels, the degree of association between dietary fat assessed by a semiquantitative food frequency questionnaire and blood lipid levels observed in this analysis may not apply to all other populations. In addition to adiposity, physical activity, and age, insulin resistance is also affected by genetic predisposition and other dietary factors. Furthermore, the higher level of education and motivation in this cohort of health professionals may have improved accuracy in completing the questionnaire. Further, a wide range of dietary fat exists within this population, which facilitates detecting associations. Limitations of this study include a large number of men who were excluded and a modest sample size. However, these exclusions are unlikely to affect the validity of the associations investigated because they were largely for missing data on variables not used in this analysis, and the replication of the results in our parallel cohort of women adds confidence in this finding.

As predicted by controlled feeding studies, in a cross-sectional analysis we found that higher dietary total fat intake measured by a self-administered semiquantitative food frequency questionnaire was strongly associated with lower plasma fasting triglyceride levels and more weakly associated with higher HDL cholesterol levels. Along with the nearly identical results for women in the Nurses’ Health Study, these data do not support the suggestion that the validity of dietary fat assessment by our self-administered questionnaire has been seriously overestimated by comparisons with weighed diet records due to correlation of errors between methods (14–16, 27). These findings illustrate the value of biologic markers for assessing the validity of dietary questionnaires if they are sensitive to the dietary factor being evaluated, even if they are not specific for dietary influences. For assessing the validity of total fat intake, fasting triglyceride levels appear to be more sensitive than HDL cholesterol, although they may be less useful indicators of dietary fat in groups that do not include overweight persons. Our findings provide additional evidence that informative measurements of dietary fat can be obtained by carefully constructed food frequency questionnaires and that important associations between intake of total fat and disease risk can be detected by such methods.

**ACKNOWLEDGMENTS**

Supported by grants CA55075, HL35464, and AA11181 from the National Institutes of Health.

The authors thank Al Wing for programming assistance and Christina Mathews for manuscript preparation.

**REFERENCES**


Am J Epidemiol Vol. 154, No. 12, 2001