In epidemiologic investigations of diet and health outcomes, food-frequency questionnaires (FFQs) are often used in large cohorts to characterize participants’ average food intake. Central to the application of FFQs in these studies are comprehensive food-composition databases that document the specific nutrient content of various foods. However, myriad other factors, such as physical form and particle size, are not typically included in food-composition databases but influence the in vivo biological effects of foods consumed (1). Thus, it is important to classify foods according to biological effects obtained directly from metabolic experiments to gain further insights beyond those determined by chemical analysis alone. For carbohydrate-containing foods, a large body of experimental evidence has now accumulated regarding their abilities to increase blood glucose (1, 2). In 1981, Jenkins et al (3) developed the concept of glycemic index (GI), which ranks how much blood glucose increases after ingestion of a particular carbohydrate-containing food relative to that of a standardized source (eg, pure glucose or white bread). Although the clinical utility of GI in the dietary management of diabetes (hyperglycemia and hyperlipidemia in particular) has been documented (4), concerns were repeatedly raised about the potential for rating foods “good” or “bad” solely on the basis of their GI values. Because the amount of carbohydrate in a food is the major determinant for blood glucose response, dietary glycemic load (GL) was subsequently introduced to quantify the total glucose-increasing potential of carbohydrate-containing foods (1). In 1997, Salmeron et al (5, 6) reported the first 2 prospective studies directly linking dietary GI and GL to increased diabetes risk in 2 cohorts of men and women followed for 6 y. Since then, more than a dozen prospective studies (7–17) have reported their findings, revealing a generally positive—albeit heterogeneous—trend relating dietary GI and GL to diabetes risk (Figure 1).

In this issue of the Journal, Sluijs et al (18) contribute additional observations in support of this positive trend from 37,846 men and women aged 21–70 y followed for 10 y. Analyzing data from the Dutch component of the European Prospective Investigation into Cancer and Nutrition cohort (EPIC-NL), the authors reported that increased dietary GI and GL were significantly associated with increased diabetes risk. The findings for different effects of simple compared with complex types of carbohydrates on diabetes risk appeared to be mixed in the EPIC-NL. Interestingly, dietary carbohydrate and sugar intakes were inversely related to diabetes risk in the univariate analysis, but after further adjustment for energy intake and lifestyle factors, a positive association with diabetes risk was observed, especially for intake of starch. Sluijs et al further excluded those who underreported energy intake and observed that the GI–diabetes relation was strengthened in their sensitivity analysis. Measurement error can occur in dietary GL just as it can in any aspect of diet in free-living humans. Such errors in assessing diet GL, however, are likely to be unrelated to the outcomes of interest in a prospective setting. Thus, those previous studies with a small sample size and/or lack of rigor in dietary assessment and follow-up may have underestimated the underlying association between dietary GL and diabetes risk by reporting null results.

Aside from errors due to dietary assessment, differences in specific statistical models used in different studies, although less well appreciated, may have also accounted for the heterogeneous findings in the literature. Dietary GL is a function of total carbohydrate intake and GI values for foods. It is, by definition, highly correlated with intake of carbohydrates and total energy. To illustrate the various adjustment for variations due to individual differences in total energy and carbohydrate intake, we herein describe Models A–C. In the formulas below, a pooled logistic model is used with emphasis on macronutrient composition to simplify discussion. Y represents diabetes/outcome status; Pr(Y = 1|X_i) represents the probability of having diabetes, given {X_i}, the exposures of interest, and other covariates including alcohol and other dietary factors.

(Model A) \( \logit \{p r(Y = 1|X)\} = a_0 + a_1 \cdot GL + a_2 \cdot Protein + a_3 \cdot Total \ energy \ intake + \Sigma a_i \cdot Other \ covariates \)

In Model A, \( a_1 \) represents both the effect of increasing quantity of carbohydrates and the effect of high GI (quality of carbohydrates). Because fats are not included in this model, the effect due to increasing quantity of carbohydrates can only be achieved by reducing intake of fats (ie, substituting carbohydrates for...

1 From the Program on Genomics and Nutrition, Center for Metabolic Disease Prevention, Departments of Epidemiology and Medicine, UCLA, Los Angeles, CA.
2 SL is funded by the Burroughs Wellcome Fund, DK062290, DK062290-06A2, DK066401, DK084452, and U01-HG-005152.
3 Address correspondence to S Liu, UCLA Center for Metabolic Disease Prevention, 650 Charles E Young Drive South, Box 951772, Los Angeles, CA 90095-1772. E-mail: siminliu@ucla.edu.
First published online September 1, 2010; doi: 10.3945/ajcn.2010.30187.
One can further examine the glycemic effect of carbohydrates by including fats in the model:

\[
\text{Model B} \quad \text{Logit} \{ pr(Y = 1 | X_i) \} = \beta_0 + \beta_1 \cdot \text{GL} + \beta_2 \cdot \text{Protein} + \beta_3 \cdot \text{Fats} + \beta_4 \cdot \text{Total energy intake} + \Sigma \beta \cdot \text{Other covariates}
\]

In Model B, \( \beta_1 \) represents the glycemic effect of total carbohydrate. Because intake of protein, fats, and total energy are included simultaneously with GL in this model, \( \beta_1 \) can be interpreted as the effect of substituting foods with high GI for foods with low GI while holding constant the intake of protein, fats, and total energy.

For the purpose of separating the glycemic effect of carbohydrate on diabetes risk from its general nonglycemic effect, one can also consider the following alternative model:

\[
\text{Model C} \quad \text{Logit} \{ pr(Y = 1 | X_i) \} = \beta'_0 + \beta'_1 \cdot \text{GL} + \beta'_2 \cdot \text{Protein} + \beta'_3 \cdot \text{Carbohydrate} + \beta'_4 \cdot \text{Total energy intake} + \Sigma \beta'_1 \cdot \text{Other covariates}
\]

In Model C, \( \beta'_1 \) can still theoretically be interpreted as the effect of substituting high-GI carbohydrates for low-GI carbohydrates.

**FIGURE 1.** A: Estimated relative risk of type 2 diabetes comparing the highest and lowest categories of dietary glycemic load as reported in 13 prospective studies of men and women weighted by person-years. B: Cumulative meta-analysis of estimated relative risks of type 2 diabetes in 13 prospective studies comparing the highest and lowest categories of dietary glycemic load weighted by person-years. The highest and lowest *quintile, **quartile, tertile, and decile values were compared. *Per 10 units; †Per increase in SD. The length of the horizontal lines indicates the width of the 95% CIs.
However, the high correlation between GL and carbohydrate intake creates the need for careful analysis and thoughtful interpretation, especially when both carbohydrate intake and GL are further categorized, leaving room for residual confounding. In this regard, the wide variation in dietary GL compared by different models in previous studies also contributed to significant heterogeneity in findings (Figure 1).

Interpreting properly the biological meanings of the specific regression coefficients from these models is important because they imply different preventive strategies that could be implemented in intervention trials. Sluijs et al (18) indicated that their findings for the GL-diabetes relation similar to Model A were most robust. It is indeed the only model in which the term GL captures the effects on diabetes risk due to the quality and quantity of carbohydrate intake and their interaction. While GL does represent quantitatively the glycemic equivalent of carbohydrates standardized to a reference source (used in determining the GI), adding the term for fats or carbohydrate, the coefficient for GL would represent the effect from the quality of carbohydrate intake (ie, the effect of dietary GI weighted by the amount of carbohydrate consumed) (2). From the perspective of preventive measures, a low-GL diet can have several different macronutrient compositions with varying percentages of energy from carbohydrates, fats, and proteins.

In nutritional epidemiology, the importance of adjusting for total energy intake has long been recognized, and various statistical models for the adjustment of total energy intake have been developed and compared (19). Because total energy intake is the sum of carbohydrate, protein, and fat intake, isolating the specific effect of one macronutrient from the others and from the effect of total energy on disease risk is statistically impossible. However, this issue of statistical colinearity merely reflects a biological reality and demands the application of data from metabolic experiments to go beyond that of the food-composition table alone. Because hyperglycemia and hyperinsulinemia play critical roles in the pathogenesis of type 2 diabetes and related complications, the biological effect of different foods on blood glucose, insulin, and other intermediaries should be carefully determined and quantified. By incorporating GI values of foods into dietary assessment of GL, we have gained significant insights beyond simply investigating the relation of carbohydrates (compared with fats) with diabetes risk. These strategies should be extended to the investigation of other highly correlated dietary variables with disease outcomes whenever possible. Statistical models building should also be guided with clear specification of strategies for intervention. Such standardization will also enable future cumulative meta-analyses (Figure 1B) to obtain more robust estimates relating dietary factors to health outcomes in human populations.

Neither author declared a conflict of interest.

REFERENCES