

Errors in the interpretation of dietary assessments¹⁻³

George H Beaton, Jan Burema, and Cheryl Ritenbaugh

ABSTRACT Two years ago, I reviewed the analytic effect of error in the estimation of dietary intake, describing the emphasis on the "random" day-to-day variation in reported intake. Interest in this area is increasing and there are signs of progress in analytic strategies. This paper focuses on two concerns about the use of dietary data in analyses. The effect of different methods of adjusting analyses of fat and a health outcome for energy is illustrated through an exploration of the association between fat intake and high body mass index in data sets from the US Department of Agriculture and a Dutch national survey. Both a shift in the analytic question and a change in the error structure occur as analysis strategies are changed, leading to confusion in interpretation. The paper also addresses the growing concern about bias in the estimation of intake and the possibility that differential bias moves with stratification variables of analytic interest. The increasing use of doubly labeled water estimates of energy expenditure as a gold standard for checking on overall bias in reporting is commendable. There will always be error in dietary assessments. The challenge is to understand, estimate, and make use of the error structure during analysis. *Am J Clin Nutr* 1997; 65(suppl):1100S-7S.

KEY WORDS Measurement error, dietary assessment, distribution adjustment, bias, adjustment for energy

INTRODUCTION

Two years ago, at the First International Conference on Dietary Assessment Methods, I presented a paper on the relation between planned analyses and the choice of dietary assessment method (1). This paper addresses a different dimension of that issue.

The earlier paper contained two messages. The first was that there is not, and probably never will be, a method that can estimate dietary intake without error. The second was that different types of error have different effects in analysis and interpretation. The conclusion was that the real challenge is to develop methods for assessing the error structure of dietary data sets and statistical approaches to analysis that will better take that error structure into account. I did not argue simply for validation studies (which are desperately needed) but also for using such studies to characterize and describe the nature and magnitude of error terms. This was a rather unorthodox way of thinking about "validation."

I now take up the same story, but not in a review of progress. Instead, I cite some interesting analytic developments; illustrate an interpretation conundrum that surfaced a while ago, is

not yet resolved, and may be important for many of us; and finally, backtrack ≥ 20 y in suggesting that what has become extremely obvious in the past 5-10 y is something that dietary investigators were worrying about 20-30 y ago—something that was swept under the table by people like me when interest in day-to-day variation and random error emerged in the 1980s: the problem of bias in the estimation of intake.

SOME MARKERS OF PROGRESS

In 1992 I expressed the hope that there would be an increased effort to estimate error terms and use them in analyses (1). Willett et al (2), following methods described earlier (3), included such adjustments in their analysis of the relation between fat intake and breast cancer. More recently, Schmid and Rosner (4) described a Bayesian approach to logistic regression in the presence of measurement error. Turning to the estimation of error itself, Bhargava and Bouis (5) used a random-effects model in maximum-likelihood analyses to examine the partitioning of variance in children in the Philippines, and then, in a Houston population, to assess the effect of a nutritional intervention on patterns of variation in the consumption of target nutrients (6). Plummer and Clayton (7, 8) adopted a new approach to validation of dietary assessment methods that uses covariance analyses and Kaaks et al (9) proposed still another method.

The methods described by these authors may not represent final or even widely adopted solutions. Nevertheless, they mark a growing interest in error terms and, more important, in applying new approaches to old issues. This sort of innovation is probably necessary for important progress. At this meeting we heard about more examples of augmented interest in the validation of dietary assessment methods and the estimation and use of error terms in analyses. Certainly the titles in the program suggest that a major evolution of thinking is under way.

¹ From the Department of Nutritional Sciences, Faculty of Medicine, University of Toronto; the Department of Human Nutrition, Wageningen Agricultural University, Wageningen, Netherlands; and the Department of Family and Community Medicine, University of Arizona, Tucson.

² Keynote address by G Beaton; J Burema and C Ritenbaugh contributed in important ways to the concepts presented and provided access to data used in some of the examples included.

³ Address reprint requests to GH Beaton, Department of Nutritional Sciences, University of Toronto, Toronto, Ontario, Canada M5S 3E2. E-mail: g.beaton@utoronto.ca.

My personal interest in dietary intake data relates primarily to the assessment of the apparent nutrient adequacy of observed intakes. For this application, I am interested in distributions of "usual intakes." At the meeting 2 y ago, we heard preliminary reports from two groups developing methods to adjust observed distributions and better estimate the distribution of usual intakes (ie, remove the effect of day-to-day variation) (10, 11). One of these methods, the semiparametric Iowa State University (ISU) technique (10), has continued to develop (12) and is now available as a SAS (SAS Institute Inc, Cary, NC) program in beta testing (KW Dodd, AL Carriquiry, WA Fuller, *Estimation of usual dietary intakes: a user's guide to ISU software*, unpublished, 1994). This technique can be used to make the necessary adjustments as long as there are at least two observations per person. It appears to have a considerable advantage, particularly for highly skewed distributions, over the simple adjustment method proposed in a National Academy of Sciences (NAS) report (13).

A comparison of the results achieved with the two methods in an estimation of distributions of vitamin C intakes is shown in **Figure 1**. Two days of data were selected at random from a data set of 600 nonpregnant, nonlactating women, each of

whom contributed 6 d of data. The original data had been extracted from the US Department of Agriculture (USDA) 1985 Continuing Survey of Food Intakes by Individuals (CSFII-85) by Cheryl Ritenbaugh and colleagues for another project. The graphs show the distribution of the 2-d and 6-d means and the estimated distributions of usual intakes derived with the ISU program and with the NAS method, both using only the randomly selected 2 d of data. Although the two derived distributions appear similar, the ISU derivation is on much stronger theoretical grounds (12).

The ISU program has several additional desirable features. It takes into account sampling weights from complex survey designs and will output adjusted intake distributions, percentile distributions, and probability densities above or below input cutoff values. Unfortunately, it takes a long time to run on most personal computers.

The importance of adjustments in nutritional evaluations of observed intakes can be illustrated by looking at the Dutch national survey data set used to examine the apparent adequacy of vitamin A intakes in women. The survey collected 2 d of data and information was available for 1185 nonpregnant, nonlactating women (14). The results of applying the proba-

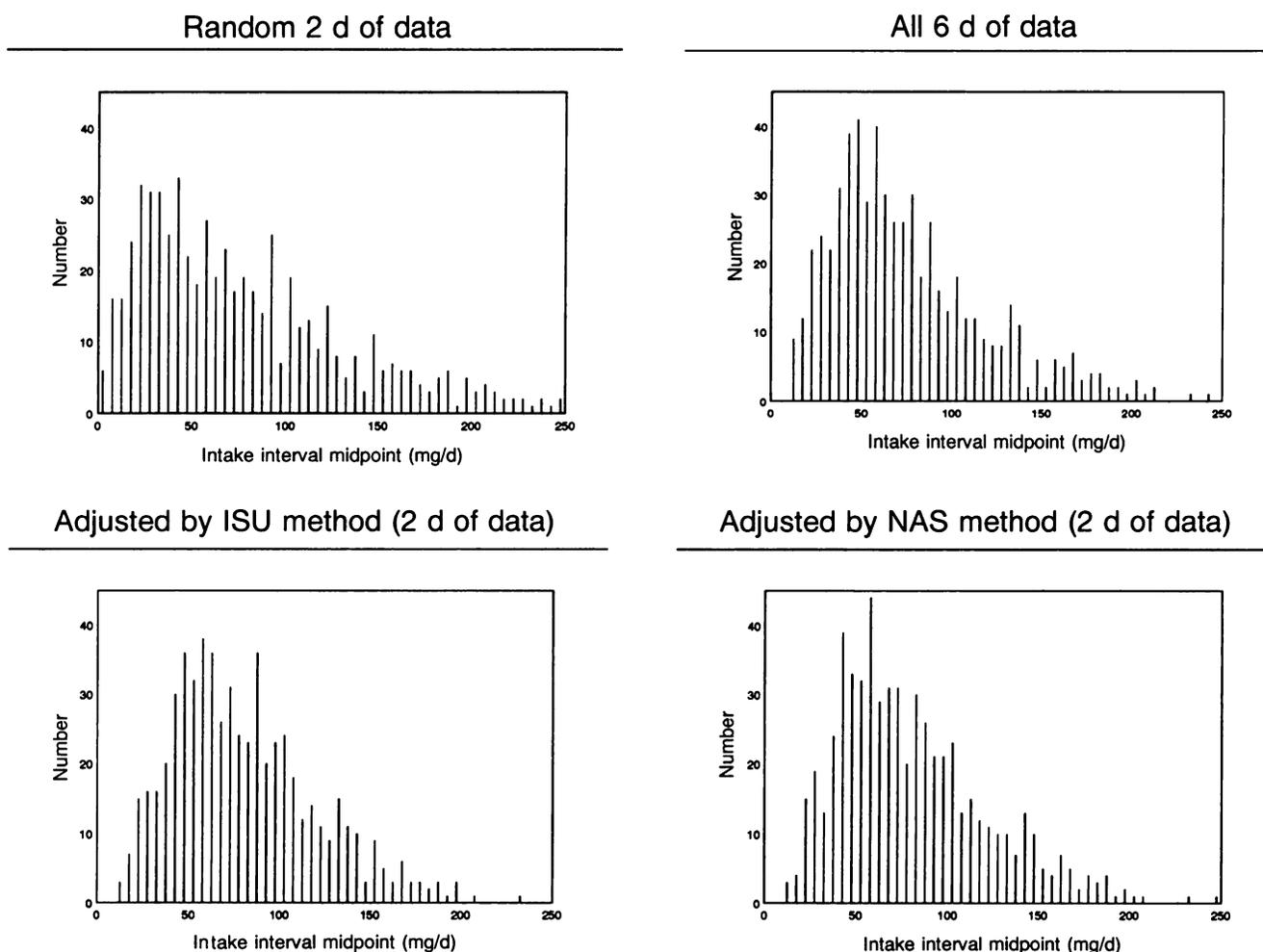


FIGURE 1. Comparison of derived distributions of vitamin C intake. A 2-d sample of data drawn from the US Department of Agriculture 1985 Continuing Survey of Food Intakes by Individuals was used for the National Academy of Science (NAS) and Iowa State University (ISU) derivations. References for the NAS and ISU methods are given in the text. The distribution of 6-d means is the best available actual description. $n = 600$.

bility approach to assessment (13), using the Food and Agriculture Organization (FAO)/World Health Organization (WHO) estimates of basal and normative vitamin A requirements (15), are shown in **Table 1**. The basal requirement estimate is the intake needed to maintain all recognized biological functions of vitamin A but that is inadequate to maintain important stores of the nutrient. The normative requirement estimate is the intake believed to be adequate to maintain stores sufficient for ≈ 3 mo; no functional advantage is attached to this requirement. The estimated prevalences of intakes inadequate to meet basal and normative percentages were calculated with use of the original (2-d) intake estimates, the original estimates adjusted with the NAS method, and the original estimates adjusted with the ISU approach. Table 1 clearly indicates that as the intake distribution is adjusted to remove the effects of day-to-day variation, the apparent prevalence of inadequate intakes decreases. In the case of the highly skewed distribution of vitamin A, even more dramatic effects would be observed in assessments for possible excessive intakes (the upper tail of the distribution).

I sincerely congratulate and thank the ISU group for continuing this work. I expect that the ISU approach will be applied in future analyses of the national survey databases and that it is probably also applicable to smaller special-purpose surveys. Further refinements of the method now being developed will expand its usefulness.

AN UNRESOLVED CONUNDRUM

Several years ago, I submitted a paper to the journal *Statistics in Medicine*. The associate editor immediately and justifiably rejected it, saying that the topic (a historical review of the interfacing of statistics and biology in interpreting and applying nutrient requirements) was not of interest to the journal's readership. He suggested that I instead address a real problem: the interfacing of biological and statistical concepts in adjusting analyses for energy intake. At the time, I thought that was not an interesting issue.

Subsequently, I read a 1992 article by Kushi et al (16) that described an epidemiologic study of dietary fat intake and breast cancer incidence conducted in 1986–1989 in about 32 000 postmenopausal women. Dietary data had been collected with a food-frequency questionnaire said to be almost identical to that used in the Nurses' Health Study by Willett et

al (17). The questionnaire had been "calibrated" against repeated 24-h recalls and apparently found comparable to the Nurses' Health Study instrument.

In their analyses, Kushi et al noted the need to control for energy intake while asking about associations between fat intake and breast cancer. In their data set, the cross-correlation between fat and energy was 0.90, implying severe confounding of any separate effects. The authors recognized that there were several ways of controlling the analyses (**Table 2**). All these methods have proponents and justifications. The residual model was proposed by Willett and Stampfer (18). The energy-partition, or energy-decomposition, approach was suggested by Howe et al (19). The two methods have been in general use for many years. There have been extensive discussions about the relative merits and interpretations of the different approaches (18–21).

What was interesting (and troublesome) was that when these approaches were applied in parallel in proportional hazards regression models, the apparent inferences about the relation between fat intake and breast cancer were different. **Table 3**, which is taken from the paper by Kushi et al (16), provides a simple comparison of the results obtained with the various approaches.

None of the adjustment methods yielded results that provided definitive evidence of an association between fat intake and breast cancer; rather, all were consistent with a weak positive association between the two. Kushi et al (16) suggested that if only one analysis method had been applied, the conclusions might have been discordant. Analysis according to the standard multivariate approach indicated that the risk of breast cancer increases (albeit not significantly) with an increase in fat intake. The analysis involving the use of residuals suggested that there is no association between fat intake and breast cancer. The results achieved with the energy-partitioning approach resembled those of the multivariate approach, whereas the findings obtained with the energy density method were similar to those of the residual technique. Kushi et al pointed out that some of the conflicting results in the literature might be traced back to the phenomenon illustrated in this table—different methods of energy adjustment.

When this possibility was brought to my attention, it was coupled with a question that I found intriguing. Kushi et al (16) addressed this question somewhat in a discussion of the public health implications of the variations in results. The question is still sufficiently interesting that, although I cannot answer it to

TABLE 1
Effect of method of adjusting intake distributions on apparent prevalence of inadequate vitamin A intakes in the Dutch national survey¹

Data set used	Level of requirement applied ²	
	Basal	Normative
		%
Original data (2 d)	2.15	16.00
Adjusted with NAS method	0.44	9.65
Adjusted with ISU method	0.04	4.19

¹ NAS, National Academy of Sciences; ISU, Iowa State University. References for the NAS and ISU methods are given in text.

² Basal and normative requirements refer to the intake necessary to meet all recognized functional needs and that required to also generate and maintain levels of storage judged to be desirable (15).

TABLE 2
Approaches to controlling for energy when examining the association of fat intake with a health outcome¹

Adjustment method	Description
Standard multivariate	Both absolute fat and energy intakes are included
Residual	Fat intake is first regressed on energy, then the residual is entered along with energy
Energy partition	Total energy is partitioned into fat and nonfat energy
Nutrient density	Fat is expressed in proportion to energy [% of energy or g/4814 kJ (g/1000 kcal)]

¹ From reference 16.

TABLE 3

Apparent relative risk for breast cancer as a function of mode of expression of fat intake¹

Adjustment method	Quartile of fat intake				Chi-square for trend (P)
	Q1 (low)	Q2	Q3	Q4 (high)	
Standard multivariate	1.00	1.17 (0.87, 1.58) ²	1.25 (0.88, 1.76)	1.38 (0.86, 2.21)	1.79 (0.18)
Fat residual	1.00	1.24 (0.93, 1.64)	1.30 (0.98, 1.72)	1.16 (0.87, 1.55)	1.11 (0.29)
Energy partition	1.00	1.15 (0.64, 1.43)	1.19 (0.87, 1.43)	1.26 (0.87, 1.84)	1.45 (0.23)
Nutrient density	1.00	1.12 (0.84, 1.48)	1.25 (0.95, 1.65)	1.13 (0.84, 1.51)	1.05 (0.31)

¹ From reference 16. Other covariates in models included age, age at menarche, age at menopause, age at first birth, family history of breast cancer, BMI at age 18 y, waist-to-hip ratio, history of benign breast disease, alcohol intake, and nonfat energy intake (for the energy-partition model) or total energy intake (for other models).

² 95% CIs in parentheses.

my satisfaction, I want to give it exposure. The question is as follows:

Are the differences in analytic results attributable to

- 1) differences in the error structure of the derived energy-adjusted variables, or
- 2) subtle but important changes in the biological question that is being asked in analysis?

To address this question, I looked at an analytic question that involved the same range of adjustments for energy as that described by Kushi et al (16) and thus might also show effects of the method used to make those adjustments. The hypothesis I examined is one that most nutritionists would probably accept as gospel truth, ie, obesity is associated with high fat intake. I reasoned that if this hypothesis is true, then overweight people (a group enriched with individuals who are obese) would show higher fat intakes. I examined this assumption in the CSFII-85 data set. The particular data set used (which was prepared by Ritenbaugh for a USDA-funded study of methodologic issues in food-consumption surveys) contained information from 600 nonpregnant, nonlactating women aged 18–55 y who had completed all six waves of data collection.

The subjects were categorized as thin [body mass index (BMI, in kg/m²) < 18.5], normal (BMI 18.5–25), or overweight (BMI > 25) with use of one of the several classification schemes. The first examination looked at the mean fat and energy intakes in the three groups. Residuals of the regression of fat on energy, decomposed energy intakes, and nutrient densities were then computed and group means compared. The results obtained are shown in **Table 4**.

By analysis of variance, intakes of energy and nonfat energy, but not of fat or fat energy, were associated with weight status. However, lower values for each of these were observed in the high-BMI (overweight) group. Total fat intake, fat as a percentage of energy, and fat residual were not demonstrably associated with weight status, although the results did suggest a relation between fat residual and weight status.

The initial conclusion to be drawn from these results seems to be that the hypothesis was wrong—a high fat intake is not associated with obesity. I will return to this issue later, because there are other obvious explanations for this finding, and will instead focus on the different impressions gained when the derived values are used.

A substantial body of evidence indicates that total food intake was underestimated in the USDA CSFII-85 survey (a topic to which I will return). Therefore, to determine whether the findings shown in Table 4 resulted from peculiarities in that data set, I repeated the same analyses on another data set. In cooperation with Jan Burema and Wija van Staveren of Wageningen Agricultural University, Netherlands, I analyzed the Dutch national survey. This survey collected 2 d of data. The data set I used included information from 1165 nonpregnant, nonlactating women aged 20–50 y. Subjects reporting special diets had been excluded. The results of these analyses are shown in **Table 5**. Interestingly, energy intakes in the Dutch data set are ≈2500 kJ/d higher than the estimated intakes in the USDA data set in each of the categories of weight status.

TABLE 4

Summary statistics from US Department of Agriculture survey data, with subjects classified according to BMI status¹

	BMI status		
	Thin (n = 26)	Normal (n = 376)	Overweight (n = 198)
BMI (kg/m ²)	17.7 ± 0.1 (16.0–18.4) ²	21.8 ± 0.1 (18.5–24.9)	30.9 ± 0.4 (25.1–52.3)
Energy (kJ/d)	6873 ± 394	6434 ± 102	6007 ± 124
Fat (g/d)	66.4 ± 4.1	63.3 ± 1.2	59.1 ± 1.5
Fat adjusted for energy ³	62.0 ± 0.5	62.3 ± 0.7	60.5 ± 1.8
Fat energy (kJ/d)	2503 ± 155	2382 ± 43	2227 ± 56
Nonfat energy (kJ/d)	4374 ± 285	4052 ± 63	3780 ± 75
Fat (% of energy)	36.5 ± 1.5	36.1 ± 0.3	36.1 ± 0.4
Fat residual	–1.70 ± 2.63	–0.08 ± 0.45	0.38 ± 0.60

¹ $\bar{x} \pm SE$.

² Range in parentheses.

³ Adjusted fat intakes derived from covariance analysis with energy intake and weight status as covariates.

TABLE 5

Summary statistics for Dutch national survey, with subjects classified according to weight status¹

	Weight status		
	Thin (n = 52)	Normal (n = 915)	Overweight (n = 218)
BMI (kg/m ²)	17.9 ± 0.1 (15.2–18.5) ²	21.8 ± 0.1 (18.5–25.0)	28.0 ± 0.2 (25.0–41.8)
Energy (kJ/d)	9825 ± 354 ^a (4575–17 292)	9071 ± 77 ^b (3278–23 856)	8665 ± 146 ^b (4048–17 024)
Fat (g/d)	109.5 ± 4.9 ^a	99.7 ± 1.1 ^b	99.1 ± 2.2 ^b
Fat adjusted for energy ³	99.5 ± 2.2 ^{a,b}	99.1 ± 0.5 ^a	103.7 ± 1.1 ^b
Fat energy (kJ/d)	4123 ± 184 ^a	3755 ± 41 ^b	3734 ± 82 ^b
Nonfat energy (kJ/d)	5701 ± 212 ^a	5320 ± 46 ^b	4931 ± 85 ^c
Fat (% of energy)	41.3 ± 0.9	40.8 ± 0.2	42.4 ± 0.5
Fat residual	-0.64 ± 2.34 ^a	-0.89 ± 0.52 ^a	3.80 ± 1.17 ^b

¹ $\bar{x} \pm SE$, by Duncan's test. Means in the same row with different superscript letters are significantly different, $P = 0.05$.

² Range in parentheses.

³ Adjusted fat intakes derived from covariance analysis with energy intake and weight status as covariates.

As in the USDA data, total fat intake and energy intake were lowest in the high-BMI group of subjects. The negative association was pronounced with respect to nonfat energy. However, the levels of fat adjusted for energy by covariance analysis and of fat residual were significantly higher in the overweight group compared with the thin group. The level of fat residual was higher than in the normal-weight group. Yet, fat as a percentage of energy was similar in all three groups.

What is common to both data-set analyses is that the apparent answer to the original question differs according to how one controls for energy. In both the USDA and the Dutch analyses, fat intake was not higher in the overweight group. Rather, in both data sets, energy intake, particularly nonfat energy intake, was apparently lower in the high-weight group than in either of the others. Fat energy, of course, is the same as fat. The USDA data suggest a positive association between fat residual and overweight, an association clearly seen in the Dutch data. An association cannot be established in either study, however, when fat is expressed as a proportion of energy.

In these analyses, differences in the random-error terms of the derived variables are unlikely to have great importance in determining the apparent direction of an association. The data are categorized according to BMI, and misclassification because of random error is unlikely to be a serious problem. The presence of random error, however, would increase the SE and thereby reduce the statistical power in tests of significance. If error is not the principal reason for the differences in apparent results in the derived variables (Tables 4 and 5), then we are left with the likelihood that all the results are valid but that different biological questions were being asked.

One can readily appreciate that there is an important difference between asking whether it is the total fat intake or the fat concentration that affects an outcome. Biologically, we might be talking about differences in metabolic loading or variations in the milieu of the gut. It is more difficult to distinguish between asking about the concentration of fat in the diet and the variation in fat intake after taking energy intake (the fat residual) into account. These are similar. Conversely, looking at the fat and nonfat components of the diet as distinct entities seems to imply something different—but perhaps not that different from the implied question in multivariate analyses that include both fat and energy as covariates.

What is interesting in both the USDA and Dutch data is that the nonfat energy intake appears to have a much clearer asso-

ciation with weight status than do any of the other variables. Does this mean that much or most of the apparent association between energy intake and weight status really involves the nonfat energy component? Again, Kushi et al (16) carried these considerations a step further when they discussed a public health interpretation of analytic results. If total fat intake was clearly related to an undesirable outcome, one might feel comfortable advocating a reduction in fat intake. But what should be recommended if total fat intake is not so related, but the fat nutrient density or fat residual is? This could imply that a person with a low fat intake but a high nutrient density (ie, a low energy intake) is at as great, or perhaps greater, risk than one with a high fat intake but a lower nutrient density (a high energy intake). If we assume that activity is driving the high energy intake, such a scenario is feasible. We would, however, still have difficulty translating the results of the analysis into dietary advice.

If my original question is asked in a cause-and-effect manner—does fat intake predict overweight?—the fat measure becomes the independent variable and random error might then have an appreciable effect: the higher the within-person variation relative to the between-subject differences, the more likely a bias in the results of regression or classification analyses (1). The results of a partitioning of variance for these same variables (omitting the adjusted fat intake) are shown in Table 6.

Obviously, fat and fat energy have an identical error structure, because fat energy is simply fat \times 9. Interestingly, nonfat energy appears to have a much more favorable variance ratio (the best of all the variables). Not surprisingly, when the between-person differences in total energy intake are removed, either by computing the nutrient density or applying the regression residual approach, the variance ratio is increased (between-person variation reduced with less change in within-person variation). The correlation between the fat residual and fat as a percentage of energy was 0.85 for the USDA data and 0.92 for the Dutch data. It can be inferred that in conventional epidemiologic analyses of a question such as, does fat intake predict a high body mass index? fat or fat energy should be less problematic variables than either fat residual or fat nutrient density.

There is, however, another important distinction between them. Fat residual is essentially independent of energy intake, whereas fat as a percentage of energy must be correlated with



TABLE 6
Variance ratios in US Department of Agriculture (USDA) and Dutch data sets

	USDA			Dutch		
	SD _{inter}	SD _{intra}	V _{ratio}	SD _{inter}	SD _{intra}	V _{ratio}
Fat	18.4	28.7	2.4	23.4	32.7	2.0
Fat energy	165.7	258.7	2.4	210.9	294.6	2.0
Nonfat energy	249.9	336.2	1.8	258.5	283.3	1.2
Fat (% of energy)	4.4	9.0	4.2	4.8	6.8	2.0
Fat residual	6.7	14.5	4.8	11.3	16.0	2.0

energy [in the Dutch data, the correlations between energy intake and fat residual and energy intake and fat as a percentage of energy were -0.04 and 0.21, respectively, and more surprising, the residuals were strongly correlated (-0.48) with nonfat energy]. If both energy and the derived variable are used in analysis, the nutrient density variable would still present problems, although perhaps not as serious as those introduced when energy and a component of energy, fat, are both included or fat density and energy are both entered in an analysis model. Because fat energy and nonfat energy are correlated (0.58 in the Dutch data), they cannot be treated as independent variables, even though it may seem intuitively correct to do so.

Flegal and Larkin (22, 23) estimated the relative errors in a food-frequency instrument and concluded that because of differences in error structure between nutrient vectors estimated by the same instrument, the use of energy adjustments might involve potentially serious errors.

In light of these analyses and the studies of fat intake and breast cancer mentioned earlier, we are left with two conclusions. First, there is a difference in the question being asked and that difference is likely to be an important part of the apparent discrepancy in results. Second, it is likely that the error structure differs among the variables and that this too can influence results. If it really is a measure of fat relative to energy that is important to the disease outcome, we must ask whether we should examine the nonfat energy component more closely.

These examinations of associations between weight status and fat intake are not adequate even to begin to address the causal question. We want to control the analyses for many other confounders and, above all, to examine a prospective study, thereby avoiding the real possibility that overweight

people tend to control fat and energy intakes. If that is the case, then lower intakes of both would be expected in examinations of the type I have reported. This constraint on interpretation does not negate my primary purpose, which was to illustrate that the choice of fat variables is likely to change the answer one gets.

BIAS IN REPORTING

Something else came out of these analyses that is important to note. There is reason to believe that food intake is being underestimated in large US surveys (24). I therefore examined the USDA and Dutch data sets for possible underreporting bias. I calculated each subject's basal metabolic rate (BMR) with use of the regression equations published in the FAO/WHO/UNU report on energy and protein requirements (25). Estimated energy intake was then expressed as a ratio to estimated BMR. Group means were computed for each weight-status category and then compared across categories (Table 7).

The USDA data show an overall mean energy intake of ≈6500 kJ/d, whereas the Dutch data show a mean intake of ≈9000 kJ/d, about 2500 kJ/d higher. The preliminary report of the third National Health and Nutrition Examination Survey (NHANES III) on data collected in 1988-1991 suggested average energy intakes of 7325-8200 kJ/d in women aged 20-50 y (26) (means for fat as a percentage of energy were 34.0-34.9). The authors of the report suggested that there may still be underreporting of intake by women and overweight adults (26). It seems probable that there was a great deal of underreporting in the CSFII-85 survey, perhaps as much as 25-30%.

In fact, a generalized underestimation of intake clearly exists in the USDA data set. Among persons with light to moderate

TABLE 7
Basal metabolic rate (BMR) and ratio of reported intakes to BMR, according to BMI status in US Department of Agriculture (USDA) and Dutch national surveys¹

	BMI status		
	Thin (n = 23 USDA, 52 Dutch)	Normal (n = 376 USDA, 915 Dutch)	Overweight (n = 198 USDA, 218 Dutch)
USDA			
BMR (kJ/d) ²	4466 ± 198 (3453-5680)	5182 ± 39 (3202-6371)	6275 ± 55 (4240-9071)
Energy/BMR ³	1.67 ± 0.11 (0.72-2.52)	1.28 ± 0.02 (0.39-3.71)	0.97 ± 0.02 (0.21-2.05)
Dutch			
BMR (kJ/d) ²	5316 ± 34 (4680-6099)	5768 ± 11 (4663-7104)	6379 ± 26 (5626-8510)
Energy/BMR ³	1.85 ± 0.07 (0.85-3.35)	1.58 ± 0.01 (0.58-4.28)	1.37 ± 0.02 (0.62-2.85)

¹ $\bar{x} \pm SE$; range in parentheses.

² BMR computed with FAO/WHO/UNU equations (25) with weight, height, and age.

³ By ANOVA, the three ratios were different from each other in each of the surveys.

activity levels, the expected ratios of energy intake to BMR are ≈ 1.4 – 1.6 , ie, the ratios in the Dutch data. However, the USDA data yield ratios well below this except in the thin group.

In addition, in both data sets, the ratio falls with an increase in BMI. Part of this decrease may be an artifact. The equations for estimating BMR may not adequately account for body composition, which differs across BMI groups. It is possible that BMR is underestimated in the thin group and overestimated in the overweight group. It is also likely that some of the apparent underreporting in the overweight group reflects truly low intakes (because of little physical activity, dieting, or both). Ballard-Barbash and colleagues (personal communication, 1994) have convincingly argued that the low intakes among overweight women (BMI > 27.3) in the CSFII-85 database are due partly to an increased rate of dieting.

The point is to recognize that there is considerable underreporting of intake in the USDA data and that it may be greater in the high-BMI group. This raises a question that cannot be directly answered: what was the composition of the unreported intake? If, for example, fat sources were underestimated more than other macronutrient sources, that might explain why total fat intake, even when adjusted for energy intake, did not appear to relate to weight status. Might it also explain why the analyses of weight status and intake seem to give clearer outcomes in the Dutch data than in the USDA data? Without direct information about what is not there—information we cannot get—these questions are only rhetorical.

The warning, however, is not rhetorical. If known or suspected bias in reporting exists, the macronutrient and energy intakes are likely to be wrong. If we have reason to believe that everything is underestimated proportionately, some of the energy-adjustment procedures should take that into account, and they would then be preferred in analyses. If there are differential biases in the reporting of macronutrient intakes, however, energy-adjustment procedures would not solve the problem. Indeed, in some circumstances, they would aggravate it.

As mentioned earlier, Mertz et al (24) challenged the large survey databases. Many other authors have issued similar challenges with regard to other study data (27–31). Since the doubly labeled water method for estimating total energy expenditure became available, several validation studies have been completed (32–40). The results are not encouraging. Most of the reported studies indicate a considerable underestimation of intake. Furthermore, it is not certain that this resulted from the method of data collection (repeated 24-h recall, food record, or food frequency and diet history). It may reflect either particular populations or a particular investigator's implementation of standard methods.

This has many implications for those who attended this conference. First, until more and better validation studies are performed, the fact of underestimation challenges the credibility of many of the investigations reported. Second, it casts doubt on the specific recommendation of the 1992 conference concerning calibration of food-frequency questionnaires against repeated 24-h recalls (41). That recommendation assumed that even if not unbiased, the 24-h recall constituted a reproducible method and that results obtained with it would be comparable across centers and cultures. Recent doubly labeled water studies have challenged that assumption. I reviewed the results of a validation study conducted in a setting in which the subjects were highly cooperative and the interviewers were

highly trained and used standardized procedures. Underreporting > 25% was found. Other studies with apparently less well-trained interviewers and less cooperative subjects have used the same 24-h recall with apparently less underreporting.

We now have a gold standard for the estimation of total food intake as energy. Also, broader application of that method will lead us into a new era of self-flagellation. We will begin to discover errors that we thought and hoped had been eliminated. The good thing about this method is that there is no reason why it cannot be applied to the validation of food-frequency and diet-history techniques just as it has been applied to dietary records and 24-h recalls. The door is open for a new voyage of discovery—and the challenge for the next international conference is clear.

CONCLUSIONS

In both 1992 and today, I argued that we really are moving ahead. We are gaining better insights into the nature of error in dietary data and gradually developing statistical methods to take this error into account. Problems certainly remain, but progress is being made. I also suggested strongly that we must reexamine the issue of bias in the collection of dietary data. In doing this, it will be important to test whether bias moves with the major stratification variables of interest.

We must continue to worry about the biological interpretation of statistical analyses. This will become even more important—and difficult—as we move into increasingly complex analysis strategies to deal with error terms. Obviously, as many have said before, continuing contact and dialogue are necessary among those interested in the collection and use of dietary data and statisticians knowledgeable about techniques for dealing with error in measurements. Conferences like this one constitute steps in the right direction toward this type of communication.

To end on a positive note, I remain confident that, with continued work, we will see the day when apparent conflicts in epidemiologic analyses will begin to be resolved. A gradual congruence of evidence will then occur, along with the development of a much greater confidence in the advice about diet and health that we offer to the public. 

REFERENCES

1. Beaton GH. Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology. *Am J Clin Nutr* 1994;59(suppl):253S–61S.
2. Willett WC, Hunter DJ, Stampfer MJ. Dietary fat and fiber in relation to risk of breast cancer. An 8-year follow-up. *JAMA* 1992; 268:2037–44.
3. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *Am J Epidemiol* 1990;132:734–45.
4. Schmid CH, Rosner B. A Bayesian approach to logistic regression models having measurement error following a mixture distribution. *Stat Med* 1993;12:1141–53.
5. Bhargava A, Bouis H. Maximum likelihood estimation of between and within variations in energy and protein intakes from infancy to adolescence for the Philippines. *Stat Med* 1992;11:533–45.
6. Bhargava A, Forthofer R, McPherson S, Nichaman M. Estimating the

- variations and autocorrelations in dietary intakes on weekdays and weekends. *Stat Med* 1994;13:113–26.
7. Plummer M, Clayton D. Measurement error in dietary assessment: an investigation using covariance structure models. Part I. *Stat Med* 1993;12:925–35.
 8. Plummer M, Clayton D. Measurement error in dietary assessment: an investigation using covariance structure models. Part II. *Stat Med* 1993;12:937–48.
 9. Kaaks R, Riboli E, Estève J, van Kappel AL, van Staveren WA. Estimating the accuracy of dietary questionnaire assessments: validation in terms of structural equation models. *Stat Med* 1994;13:127–42.
 10. Carriquiry AL, Jensen HH, Fuller WA, Guenther P. Methods for estimating usual intake distributions. *Am J Clin Nutr* 1994;59(suppl):305S (abstr).
 11. Aickin M, Ritenbaugh C, Woolf K, et al. A general method for adjusting dietary data to obtain stable estimates of usual intake. *Am J Clin Nutr* 1994;59(suppl):303S (abstr).
 12. Nusser SM, Carriquiry AL, Dodd KW, Fuller WA. A semiparametric transformation approach to estimating usual dietary intake distributions. *J Am Stat Assoc* (in press).
 13. National Research Council, Subcommittee on Criteria for Dietary Evaluation. Nutrient adequacy: assessment using food consumption surveys. Washington, DC: National Academy Press, 1986.
 14. Hulshof KFAM, van Staveren WA. The Dutch National Food Consumption Survey: design, methods and first results. *Food Policy* 1991;16:257–60.
 15. Food and Agriculture Organization/World Health Organization. Requirements of vitamin A, iron, folate and vitamin B₁₂. Rome: FAO, 1988. (FAO food and nutrition series 23.)
 16. Kushi LH, Sellers TA, Potter JD, et al. Dietary fat and postmenopausal breast cancer. *J Natl Cancer Inst* 1992;84:1092–9.
 17. Willett WC, Sampson L, Browne ML, et al. The use of a self-administered questionnaire to assess diet four years in the past. *Am J Epidemiol* 1988;127:188–99.
 18. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. *Am J Epidemiol* 1986;124:17–27.
 19. Howe GR, Miller AB, Jain M. Re: "Total energy intake: implications for epidemiologic analyses." *Am J Epidemiol* 1986;124:157–9.
 20. van de Vijver LPL, van der Waal MAE, Weterings KGC, van T'Veer P, Deffer JM, Kok FJ. Methoden voor energie-correctie in epidemiologisch voedingsonderzoek. (Methods for energy correction in epidemiologic nutrient research.) *Voeding* 1991;11:276–9 (in Dutch).
 21. Martin-Moreno JM. Adjustment for total caloric intake in nutritional studies: an epidemiologic perspective. *Eur J Clin Nutr* 1993;47:S51–2.
 22. Flegal KM, Larkin FA. Partitioning macronutrient intake estimates from a food frequency questionnaire. *Am J Epidemiol* 1990;131:1046–58.
 23. Flegal KM, Larkin FA, Metzner HL, Thompson FE, Guire KE. Counting calories: partitioning energy intake estimates from a food frequency questionnaire. *Am J Epidemiol* 1988;128:749–60.
 24. Mertz W, Tsui JC, Judd JT, et al. What are people really eating? The relation between energy intake derived from estimated dietary records and intake determined to maintain body weight. *Am J Clin Nutr* 1991;54:291–5.
 25. Food and Agriculture Organization/World Health Organization/United Nations University. Energy and protein requirements. World Health Organ Tech Rep Ser 1985;724:1–206.
 26. McDowell MA, Briefel RB, Alaimo K, et al. Energy and macronutrient intakes of persons ages 2 months and over in the United States: third National Health and Nutrition Examination Survey, phase 1, 1988–91. Washington, DC: US Government Printing Office, 1994. [DHHS publication (PHS) 95 1250.]
 27. Acheson KJ, Campbell IT, Edholm OG, Miller DS, Stock MJ. The measurement of food and energy intake in man—an evaluation of some techniques. *Am J Clin Nutr* 1980;33:1147–54.
 28. Livingstone MBE, Prentice AM, Coward WA, et al. Validation of estimates of energy intake by weighed dietary record and diet history in children and adolescents. *Am J Clin Nutr* 1992;56:29–35.
 29. Lichtman SW, Pisarska K, Berman ER, et al. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. *N Engl J Med* 1992;327:1893–8.
 30. Black AE, Goldberg GR, Jebb SA, Livingstone MBE, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. *Eur J Clin Nutr* 1991;45:583–99.
 31. Goldberg GR, Black AE, Jebb SA, et al. Critical evaluation of energy intake data using fundamental principles of energy physiology. 1. Derivation of cut-off limits to identify under-recording. *Eur J Clin Nutr* 1991;45:569–81.
 32. Forbes-Ewan CH, Morrissey BLL, Gregg GC, Waters DR. Use of doubly labeled water technique in soldiers training for jungle warfare. *J Appl Physiol* 1989;67:14–8.
 33. Schoeller DA. How accurate is self-reported dietary energy intake? *Nutr Rev* 1990;48:373–9.
 34. Livingstone MBE, Prentice AM, Strain JJ, et al. Accuracy of weighed dietary records in studies of diet and health. *Br Med J* 1990;300:708–12.
 35. Schoeller DA, Bandini LG, Dietz WH. Inaccuracies in self-reported intake identified by comparison with doubly labelled water method. *Can J Physiol Pharmacol* 1990;68:941–9.
 36. Bandini LG, Schoeller DA, Cyr H, Dietz WH. Validity of reported energy intake in obese and nonobese adolescents. *Am J Clin Nutr* 1990;52:421–5.
 37. Black AE, Prentice AM, Goldberg GR, et al. Measurements of total energy expenditure provide insights into the validity of dietary measurements of energy intake. *J Am Diet Assoc* 1993;93:572–9.
 38. Jones PJH, Jacobs I, Morris A, Ducharme MB. Adequacy of food rations in soldiers during an arctic exercise measured by doubly labeled water. *J Appl Physiol* 1993;75:1790–7.
 39. Davies PSW, Coward A, White A, Gregory J, Tyler H, Mills A. Total energy expenditure and energy intake in the pre-school child: a comparison. *Am J Clin Nutr* 1994;59(suppl):293S (abstr).
 40. Howat PM, Mohan R, Champagne C, Monlezun C, Wozniak P, Bray GA. Validity and reliability of reported dietary intake data. *J Am Diet Assoc* 1994;94:169–73.
 41. Buzzard IM, Sievert YA. Research priorities and recommendations for dietary assessment methodology. *Am J Clin Nutr* 1994;59(suppl):275S–80S.

