ARCHIVOS LATINOAMERICANOS DE NUTRICION
Organo Oficial de la Sociedad Latinoamericana de Nutrici6n

Vol. 45 N° 1-S, 1995

Validation of assessment methods for food intake surveys

George H. Beaton

The title of this paper can have interpretations: validation of i) food intake data collection methods, and of ii) approaches to the nutritional assessment of estimated intakes. The two topics are closely related. Believing that it was the intent of the organizers, the paper will address primarily upon the former. Unless we can reliably estimate intake, there may be limited reason to worry about how estimated intakes should be evaluated with regard to nutritional adequacy. Nevertheless, given this author's strong personal interest in the evaluation of estimated nutrient intakes, the paper will also touch on the latter.

Approaches to validation of dietary data

This field, and our understandings and approaches, are evolving rapidly. With that evolution has come the need to review and often refute former 'truths' that we had accepted. The January 1994 supplement to the American Journal of Clinical Nutrition presents the proceedings of the 1st International Conference Dietary Assessment Methods . It should provide an invaluable resource to all interested in the topic of this paper. The proceedings of a second international conference on this topic, held in 1995, is now in preparation for publication. Interested persons might also wish to read two reports from the Berlin Meetings on Nutritional Epidemiology [1].

An underlying message that comes from all of this material is:

Dietary intake cannot be estimated without error
There is no perfect method of data collection
There may be preferred methods for defined purpose

«Validation studies» for dietary data collection should not be seen as approaches to proving a method and the collected data are «valid». Rather, the real purpose of such studies should be to identify and estimate the errors present in the data so that this information can be used in analysis and interpretation.

Types of Error in Dietary Data

Error in dietary (and other) data is often classified into two broad categories: random error and systematic error or bias. This is useful since these two categories of error have rather different impacts in analysis. Note however, that the line between the two classes is very blurred. For example there are often patterns within the 'random' error' and bias in reporting may be characteristic of a particular subject but this could be randomly or systematically distributed across subjects. It is also important to recognize that in this paper, «error» is used in the statistical sense of variability and does not necessarily mean «mistake». The paper assumes that we are attempting to estimate 'usual' intakes or the average intake over extended periods of time. Deviations from this true value are treated as 'error of the estimate'.

Where do these 'errors' arise?

The first source of variation identified in Fig 1 is that which has attracted the greatest attention since about 1980 when several papers concurrently brought it to the attention of the nutrition community - the simple fact that we do not eat the same thing every day and therefore there is day-to-day variation in intake. This is inseparably mixed with the random errors that occur in reporting, recording and coding the intake of foods. We have a mixture of true variation and real mistakes. This within person 'random' error is seldom truly random. There are recurring social or other influences that give rise to patterns within this variation. Perhaps the best known pattern is that which we call the «day-of-the-week» effect. Interestingly Tarasuk and Beaton (1-2) observed that while most individuals show some sort of a seven day cycle in intakes, the cycles are not necessarily in phase across individuals and, further, that many individuals exhibited long term trends in their intake data that could not be described as «seasonal» effects. The variability of intake was very much a personal characteristic. This should not be unexpected. THinkin f the day of the week effect, if you were working in a Muslim area you might expect Fridays to be typically different. If you conduct studies in North America you could reasonably expect the non-working day to be different. This used to mean Sunday and perhaps Saturday. Today, with both husband and wife working and with 6 or 7 day a week staffing, the 'different' day off is not always on a weekend. The point I make is that these patterns, and indeed the total magnitude of the variability, are culture- and individual-dependent - and such cultural practices must be taken into account in both the design and validation of food intake methodologies.

The second major heading in the figure, «Between Person Variation», is usually the variation in which we are analytically interested - the distribution of true 'usual intakes'. There is an important, but often forgotten error component here. Individual bias in reporting (under- or over-estimation of actual intake) could be included as a random between person variable. A particular individual may consistently under- or over-report his/her intake. Across individuals the phenomenon may vary in both direction and magnitude. If this is the case it becomes extremely difficult to separate from the true between person variation. Beaton (3) suggested that random

Departamento of Nutritional Sciences. University of Toronto. Toronto-Canada.

1 Kohimeir L. and Helsing E (eds) "Epidemiology, Nutrition and Health: Proceedings 1st Berlin Meeting on Nutritional Epidemiology" Smith-Gordon, Nishimuara, Londo, 1989; and kohimeir, L (ed) "The Diet History Method: Proceedings of the 2nd Berlin Meeting on Nutritional Epidemiology", Smith-Gordon, Nishimura, London, 1991.

subjects may offer biased estimates of food intake as a differing response to food frequency instruments and that they may repeat this same error when the instrument is applied again - a subject-specific bias that is random between subjects. This would create exactly the sort of «random bias» mentioned here.

FIGURE 1
A simplified classification of error terms.

SOURCES OF VARIATION IN DIETARY DATA

- WITHIN PERSON VARIATION
  True day-to-day variation in intake
  Day-to-day error in reporting
  Either can be essentially random or patterned (e.g day of the week effect)

- BETWEEN PERSON VARIATION

  True between subject differences in their usual intake
  Systematic error (bias) in reporting but random between individuals

- BIAS
  Systematic bias in reporting - systematic across subjects
  Errors in food coding or in food composition data base.

Unwanted «error» will always be present in data sets. The approach must be to estimate what type(s) and how much and to then take this into account in analysis and interpretation.

The «random bias» is differentiated from the under- or over-reporting that is more consistent across subjects and is detected in the total data set as a systematic bias. That type of phenomenon that is much more easily detected from a comparison of group means.

In the early years of dietary validation studies, detection of systematic bias was the main focus. As interest in the distribution of intakes and in the partitioning of variance into within and between person variation increased, concern about bias was temporarily submerged. That was a serious error. Bias in estimation of intake must continue to have major priority in any nutritional assessment of intakes [(2)]. We have strong reason to believe that the large U.S. national dietary surveys (NHANES and particularly the USDA surveys) are biased toward underestimation of intake. Bias may arise during the initial reporting and recording of food intake (intentional or otherwise), from errors in the methodology of data collection (instrument errors), from errors in the food coding, or from errors in the food composition data base. We must remember also that bias in estimated intakes of population groups can arise from basic flaws in study design - from a failure to appropriately sample people and time (day of he week, season or other time-related patterns). In this case, the data may be very good in terms of error assessments - yet the data would fail to represent what the investigator or user thinks they represent. This type of error can be assessed only by looking carefully at the sampling design and at any biases that may have entered the data set because of refusals and resultant self-selection. This can become a very serious problem in surveys.

2. Beaton, G.H., Burema, J., Ritenbaugh, C. Errors in the interprettation of dietary assessments. Submitted for publication as part of the Proceedings of the 2nd International Conference on Dietary Assessment Methods, 1995.

The nightmare of the data analyst is that errors of the type discussed (random or bias) might not be independent of other variables he/she wishes to analyze. Consider some examples: eating practices often differ with cultural group. It should not be surprising that the within person and between person variances (the variance ratio) differs across cultural groups. Participation rates (and hence sampling bias) may also move with cultural groups and is often found the differ with urbanization in industrialized countries. It is at least possible that overweight individuals tend to underestimate intake while underweight persons tend to overestimate intake; if body weight status is a variable in the analysis, model then the bias in dietary data would associate with this variable and analysis and interpretation become very difficult, The only defense the analyst has is to stratify the population into high interest categories and carrying out validation tests in each stratum.

Expected Impacts of Dietary Data Errors During Analysis

As background to any discussion of validation of dietary methodology, it is important that there be a general understanding of the expected impact of error in data analysis (see, for example Liu, 1989) (4), Beaton (1994a) (5). As Fig 2 suggests, the effect differs with the type of error and nature of the analysis. It follows that depending on the perceived purpose of data collection, one type of error may be much more important than another. When it is reported that a particular data collection method has been «validated», one must immediately ask «validated for what type of use?»and this should be followed by «In what type of population?» The + and - symbols in the figure are meant to portray the relative impact of error type by analysis strategy. In looking at this or similar displays the greatest single point to recognize is that quite often the epidemiologist can live with biased data sets as long as he/she believes that the bias is uniform and that nutrients and energy are all affected proportionately. The reason is that for many of his analyses, the real need is to correctly rank/classify people whether or not their true intake has been estimated. He abhors random error because that leads to misclassification of individuals, biases regressions and generally interferes seriously with his analyses. In contrast, the nutritionist is much more likely to be interested in distributional analyses (asking about the proportion of individuals with intakes below (or above) preset cut-offs. In this case bias and random error both interfere with interpretation although the impact of modest levels of random error may be less problematic for distributional analyses than for many epidemiologic studies. This is particularly true since we now have available statistical tools that can be used to adjust distributions and greatly improve the distributional analyses. To use these new tools, the data set must include at least two repeated measures for each subject.

The reader should appreciate that nutritionists and epidemiologists have tended to approach dietary methodology selection and validation with quite different perspectives and often make seemingly conflicting choices. Nowhere is this clearer today than in the debate about food frequency/diet history methods vs multiple 24 hour recalls or food records. Each method has distinct advantages and disadvantages in terms of logistic feasibility, relative cost, and nature of the error terms that are likely to be found in the data. All too often the driving pressure for the epidemiologist is the logistical feasibility of data collection and his/her choice favours the food frequency questionnaire [often self-administered]. It is particularly difficult to estimate the errors in food frequency data.

FIGURE 2
Expected impact of various types of error on some common types of data analyses. The listing is not comprehensive and not all of the caveats have been included.

EXPECTED IMPACT OF DIETARY DATA ERROR TERMS IN DATA ANALYSIS[1]

Class of Analysis	Type of Error Systematic Bias [2]	Random Bias	Random
DIET AS INDEPENDENT VARIABLE			
Simple regression or correlation	(+) [3]	++++	++++
Categorical analyses	(+)	++++	++++
Multivariate analyses[4]	(+++)	+++++	+++++
DIET AS DEPENDENT VARIABLE			
Regression, correlation and categorical analyses	++	++	+[5]
ASSESSING ADEQUACY OF OBSERVED INTAKES			
Group prevalence	+++++	(+++)[6] +++	(+ +)[6]
Assess particular individual	++++++	+++	++++++

1 Based on Beaton (1991)
2 very serious analytical and interpretational errors arise if the bias relates to other variables of interest.
3 the intercept would be wrong in regression analyses but the slope and correlation should be unaffected.
4 in multivariate analyses, random error can attenuate or inflate detected associations depending upon the error structure (see Liu, 1989)
5 presence of random error diminishes statistical power and underestimates the correlation but it will not distort the coefficients in regression analyses.
6 There are now statistical techniques for adjustment of intake distributions to take into account the impact of these types of error but those techniques cannot be applied to the particular individual (see NRC, 1986; Nusser et al, 1995).

Approaches to Validation

What must be of great concern to both the nutritionist and epidemiologist is «How does one identify and estimate the presence and magnitude of these error terms in a data set? How does one know whether a given data set can be used for a particular proposed analysis?»

As Fig. 3 implies, there are many possible approaches in validation studies. These are not mutually exclusive. Several approaches may be included in one study. Only a few of the possible approaches are reviewed in this paper.

FIGURE 3
An overview of some common types of approaches to validation of dietary data

SOME APPROACHES TO VALIDATION OF DIETARY DATA COLLECTION METHODS

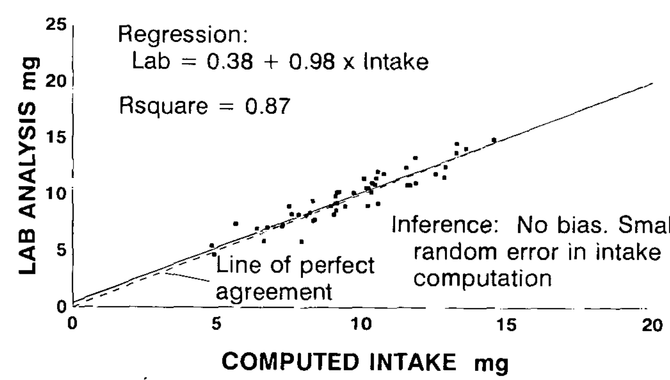
n Assessment of internal validity (food frequency instruments)
n Test-retest comparison
n Comparison with another dietary method
n Compare with an independent estimate of same reality
n Compare with biological markers of nutritional status
n Compare with chemical analysis of food composites

The preferred approach(es) varies with dietary methodology and with the purpose of collecting dietary data

Validation of the food composition data base. The most easily understood approach, is validation of the food composition data base. This approach simply involves construction of a day's food, preferably using prepared dishes collected in the field, subjecting it to laboratory analysis and then comparing computed intake with the lab results, using linear regression analysis (see Fig. 4). The intake composites should be prepared using dietary data and food samples representing different strata of the population. Such an approach serves a limited but extremely important function. It checks for a major potential source of bias and also checks to ensure that any bias detected is not related to obvious variables in the population. If discrepancies are found, then one may wish to perform analyses of particular foods. An obvious first check, before expanding food composition analyses, lies in checking the actual coding of foods. Validation of coding is usually accomplished by double coding of food records and comparison of the codes assigned. This was done as a part of a validation of the 24 hour recall methodology of the Lipid Research Clinics (LRC) trial (6). In that instance, manual coding errors were minimal and made no important contribution to the observed error terms. Conversely, close examination of coding rules (e.g. what type of fat to assume when not specified) was shown to have the potential to introduce important error.

FIGURE 4
Illustration of validation by comparison of calculated and chemically determined nutrient content of food composites. In the simulation, the lab determinations were assumed to be without error and the calculated values had a small random error component which explains both the attenuation of he regression slope and the scatter of points. If the actual error of the lab method is known, then the error of the computed value can be inferred.

SIMULATED VALIDATION OF FOOD COMPOSITION DATA BASE (FOR IRON)



50 Food composites simulated

Comparison of two dietary methods. The most widely adopted historical approach to validation involved comparison of two dietary methods applied to the same individuals. The literature abounds with such comparisons. The assumption is that this represents an independent assessment of the same reality. That assumption is not necessarily valid - a subject-specific error, or, for example, a coding error might be common to both methodologies. Nevertheless, if one can estimate the error terms in the reference method, it is possible to infer the error of the test method. In the area of quantitative recalls or records, often used as reference methods for food frequency instrument testing, one can estimate the partitioning the variance between the between-person variation in 'usual intakes' and the within person or day-to-day 'random' variation through relatively simple ANOVA procedures. This knowledge coupled with regression comparisons may be used to try to calibrate methods (Freedman et al, 1991). Much of our current understanding of the errors in dietary data arise from these 'validation studies.' Their extension to use as 'calibration tools' is more recent. Note, however, that one cannot estimate the bias component or the random bias part unless there is a third method without error or with know error structure.

Comparison of estimated intake with an external 'gold standard'. The perfect validation technique would involve comparison of intakes estimated by the test dietary methodology with intakes estimated by an independent, low error method. For many years we have compared food frequency data with multiple records of intake collected over the same time frame. Given that none of the other dietary methods is free of error, such comparisons present their own problems.

Today, it is becoming feasible, but still very expensive, to apply the doubly labelled water methodology on a sufficiently large scale to serve as a validation tool2. The assumption is that energy expenditure and energy intake should be approximately equal if measured over periods of 10-14 days in people with reasonably stable body composition. With such a gold standard, one can compare mean intake and mean expenditure to check for bias. One can also apply regression analyses to see if any bias present is consistent across all levels of intake and, at the same time, check the goodness of fit between the food intake data and the stable isotope energy expenditure estimate. From such analyses, with sufficient subjects one should be able to get a very good estimate of both the nature and magnitude of errors in the dietary method (but not necessarily where they are coming from). We have had the doubly labelled water methodology for a number of years but it is only very recently that the cost of the stable isotope preparations has started coming down to reasonable levels. Today it is the closest we can come to a true gold standard. Note that it relates only to total food (energy) intake. It does not validate nutrient intakes except indirectly.

External Validity - consistency with expectations. A much less satisfactory but still useful check on dietary intake estimates is to compute the expected Basal Metabolic Rate (BMR) using the prediction equations developed for the FAO/WHO/UNU (1985) report on energy and protein requirements. For moderately active people maintaining weight, the ratio of intake/BMR should average about 1.5-1.7. Black (7) applied such an approach, with cut-offs developed by Goldberg (8) in evaluating the results of published surveys in the United Kingdom. The result was a conclusion that «dietary assessment methods have a strong bias towards underestimation of habitual energy intake.» The same year, Mertz (9) published a paper presenting a similar conclusion about the large North American surveys. It is critically important to test intake data for possible bias, and to check to see if any bias detected moves with population strata that are of particular analytical interest.

It is important to realize that if under-reporting of energy intake is detected as a likely characteristic of the data, one cannot assume that all macro- and micro-nutrients are under-reported to the same degree. Generally thee is no way that one can be sure that the food intake not reported has composition comparable to the food use that was reported. Although many authors use energy adjustment procedures that effectively make this assumption, great care should be taken in interpretation (2).

Comparison with Biological Markers. A different approach, which is currently gaining popularity, is to compare (by regression or similar analyses) estimated nutrient intakes with the levels of biochemical markers. For example, investigators compare estimated intakes of b-carotene with serum levels of b-carotene, or compare estimated protein intake in adults with quantitative urinary nitrogen excretion. An important assumption of such approaches is that the relationship between true dietary intake and the level of the biological marker is approximately linear over the whole range of intake (or that the true relationship is known and can be applied to the data). It also assumes that between person variability in response is minimal or known and that the error of determination of the marker is very small and known. This author is personally pessimistic about the long term utility of such methods and even more concerned that an increasing number of authors suggest that we use biologic markers of intake INSTEAD of intake data.

The Test-Retest Approach to Validation. Validation approaches have become much more complex with the increased popularity of food frequency/diet history methods. Unlike the quantitative methods (24 hour recalls, diet records, etc) one cannot easily estimate the 'random error' component and hence cannot easily estimate the likely impact in data analyses. All too often we find a statement like «We used the method of Willett which has been validated...» without any further description of testing on the study population. This could be a very serious omission. The food-frequency/diet history methods are highly sensitive to cultural differences in food practices. Specific instruments cannot easily be moved across cultures and it may be very difficult to design a single instrument that is equally applicable to all cultural strata in a large multicultural study. Thus, there is a renewed need to look at validation methods specifically applicable to the food frequency instruments.

Validation testing is often based on repeated application of the same instrument to a group of subjects. Since the diet history or food frequency is usually intended to a\capture an estimate of usual intake, one would expect subjects to report very similar intakes (high correlation) in a test-retest comparison. However, if you think about it, a good correlation should arise if either (a) actual usual intakes are correctly reported each time or (b) if the same people make the same mistake each time they complete the same questionnaire. The correlation coefficient will not differentiate between these. It follows that if a test-retest design is undertaken, low correlations are interpretable - they mean that the instrument is not satisfactory. However, high correlations are difficult to interpret - they could mean that the method is very good; they could also mean that the instrument is badly flawed but individuals make the same mistakes/ misinterpretations over and over again.

Willett and colleagues have presented information about comparability of data from their food frequency instruments administered a year apart and comparisons of each of these with 28 days of dietary records collected in four one week periods during the year. This data set is portrayed schematically in Fig 4. This is probably the most ambitious validation study ever undertaken. The bottom two cells of the figure, representing the two administrations of the food frequency questionnaire include the hypothesized component of instrument error leading to a highly correlated error in the test-retest situation (see above). Beaton (3) used ANOVA and three-way correlation analyses to draw inferences about the relative error components (true between subject variation in usual intake, random error, random bias) of the food frequency instrument.

The food frequency methodology as usually implemented must be expected lose some of the real variation between people because the instrument queries only a limited list of foods and of necessity must miss some variation. At the same time, by Beaton's analyses, the food frequency data set appears to acquire additional between person variability that is very highly reproducible, presumably arising from a true error in the interface between the subject and the instrument.

Beaton (3) concluded that for analytical purposes, and depending upon the nutrient being examined, the food frequency data might be equivalent to 2 or 3 days of data collected by the quantitative methods. Willett (personal communication) noted that even if this is true, for his purposes, use of a self-administered food frequency instrument was feasible while repeated dietary recalls would be impossible. This perspective must be accepted as long as appropriate caveats are then attached to the interpretation of the analyses. However, when others report that they are using the 'Willett instrument' without reservation, trusting the assertion that it has been 'validated', they are misinterpreting the real meaning of validation as applied to dietary methods. Dr. Willett was very explicit in saying that he could live and work with the error for his planned analyses. He did not suggest that the error structure would be compatible with other types of analysis and certainly did not suggest that the instrument could be transferred to other populations without careful testing and revalidation - and many still dispute the impact of the estimated error terms in Willett's own analyses. I fear that there has been considerable naivety in the adoption and use of food frequency instruments without adequate testing. This is likely because validation is extremely difficult for these instruments.

Where do we go from here:

There is no perfect dietary methodology. There is no perfect validation method. Yet there is a need to try to estimate and understand the error structure of any data set you are using and the likely impact in analyses you perform and report.

At the 1992 International Conference on Dietary Assessment Methods, considerable concern expressed about the inadequacies of existing 'validation methods' and about the use of simple test-retest designs to validate the food frequency instruments. Some of the recommendations offered at that meeting are summarized in Fig 5. While no specific proposals were made for the perfect methodology of estimating error terms in food frequency data, one recommendation did arise. That was to apply a 'standard' method to a sub-sample of subjects as a calibration tool. In the world of optimists, this would suggest that in comparing surveys or even more, in later attempting to pool data for aggregate epidemiologic studies or meta analyses, one could compare the 24 hour recall data for consistency and also compare the reported associations between 24 hour recall and food frequency data within the individual surveys. A second International Conference on Dietary Assessment Methods was held at Harvard in January, 1995. By that time there was available a compendium of calibration/validation studies [3] as well as a resource manual on dietary methodology (10). It will be of considerable interest to see the report and recommendations that arise from the 1995 meeting. At this time, I can only leave you with a warning about limitations of existing methods and a sense of optimism that we can and will do better in the future.

FIGURE 5

Recommendations arising from the First International Conference on Dietary Assessment Methods. Based on a summary paper by Buzzard and Sievert (1994).

SOME RECOMMENDATIONS ARISING FROM THE FIRST INTERNATIONAL CONFERENCE ON DIETARY ASSESSMENT METHODS, MINNEAPLOLIS, 1992

- Encourage the inclusion of calibration sub-studies in all major surveys
- Encourage routine usage of reproducibility sub-studies (include repeated measures for at least a sample of subjects)
- Include multiple assessments of biochemical measures wherever possible and appropriate

These recommendations relate to the desirability of building in a capability to estimate error terms and to use that information in data analyses. That is the direction in which we are moving.

Nutritional Assessment of Estimated Intakes-The Probability Approach

In this section of the paper, brief reference will be made to the second aspect of dietary assessment, the nutritional assessment of estimated intake. For a more complete discussion, see (11-13).

Any approach to nutritional assessment must start with the recognition that nutrient requirements vary between individuals. At best we can describe the distribution of requirements with which the particular type of individual associates. In turn this can be converted conceptually to a description of the probability that, for a randomly selected individual his/her intake is adequate (above his/her actual, but unknown, requirement). For a group or population, this process can be repeated and averaged to obtain an estimate of the expected prevalence of inadequate intakes. A prerequisite to the application of this approach is the estimation of the distribution of usual intakes which will usually include a statistical adjustment to reduce or eliminate the effect of day-to-day variation in intake (see above).

At this moment in time, there is general acceptance of the concepts, there is also consensus that one should not assess intakes by simply comparing them to the published 'recommended intakes' or some fixed proportion of the RDA. There has been a detailed examination of the probability approach to assessment of observed intake by a National Academy of Sciences Committee (12) and there has been careful scrutiny of the statistical concepts by many other groups even though their deliberations have not been published. Conversely, there has not yet been a general implementation of the

3 Thompson, F., Moler, J.E., Freedman, L., Clifford, C., Willett, W.C. Dietary Assessment Calibration/Validation Studies Register: A Status Report, December 1994. Information available from Frances E. Thompson, National Cancer Institute, 6130 Executive Building, Bethesda, MD 20892-7344, U.S.A. Email: thompson @ dcpcepn. nci. nih. gov FAX (301) 496-9949.

approach. The plea to those of you involved in dietary assessment is «try applying the probability approach». If the distribution of nutrient requirements is approximately symmetrical (not necessarily normal), a major simplification of the probability approach yields a very similar estimate of the prevalence of inadequate intakes. One need only estimate the proportion of individuals with intakes below the mean requirement. It seems reasonable to suggest at this time that this simplified approach is applicable to all assessments except iron intake of menstruating women where wee know the requirement distribution is highly skewed.

To illustrate the approach and some of its advantages, a recent application to data arising in the Nutrition CRSP studies conducted in single communities in Egypt, Kenya and Mexico (14). Food intakes were estimated for selected individuals by monthly data collections over a period of 12 months or more. Nutrient intakes were computed using food composition tables developed for this study. These records were pooled for each individual to obtain estimates of his or her 'usual intake'. There was no further distributional adjustment made. For this presentation only the assessment of protein intake is addressed. Assessments for other nutrients have been published. Amino acid scores were computed from estimated amino acid intakes. Digestibility was estimated from an empiric relationship between dietary fibre and digestibility in mixed diets that had been studied in many laboratories. It is stressed that this was an empiric association without implication that dietary fibre affected digestibility. From these, utilizable protein intake was computed. A probability assessment was then conducted using the FAO/WHO/UNU estimates of protein requirements.

Note that the probability approach does not categorize individuals as having adequate or inadequate intakes and then count them. Everything is done as probabilities of inadequacy and prevalence is, in effect the averaged probability of inadequacy (or, if you wish, the probability that a randomly selected individual will have an inadequate intake of protein).

The results are shown in Table 1. Only in Kenya was there any suggestion that protein intake might be a problem for some children. The estimated prevalence of inadequate intakes was 2.2%. You will note also that the estimated energy intakes in Kenya were much lower than in Egypt or Mexico. Having computerized the analysis, it was a simple matter to perform 'what-if' scenarios. As shown on this slide, it was asked «What if the energy intakes had been closer to accepted estimates of requirement? Would the protein intakes have been raised above requirements?» The answer seemed to be 'yes'. This suggested that a higher priority might be given to increasing food intake than to adding protein supplements. Similarly the possible impact of amino acid fortification on protein adequacy was examined - very little benefit would be expected. In a companion papers (5, 15-16) vitamins and minerals were examined. Those papers asked about adequacy of intake for different levels of nutriture, and asked about the likely impact of changing patterns of food use - e.g. the impact of use of ascorbic acid sources and tea consumption on iron utilization. The assessment approach is flexible and encourages users to address a wider range of policy-relevant questions than is usually done.

In summary, the only real points about the probability approach to assessment of nutrient intake that are to be emphasized today are shown in Fig 7. In the judgement of this author, it is very important that those of you involved in dietary assessments take advantage of this new technology and see how it can be applied effectively in your situation.

TABLE 1
Evaluation of Protein Intakes of Toddlers
(Mean ± Standard Deviation)

Variable	Country Setting of Village Population Egypt	Kenya	Mexico
Number of toddlers included	96	100	59
Age, m*	23.6±1.6	23.8±0.9	24.8±0.9
Weight, kg*	10.9±1.2	10.2±1.2	10.8±1.1
Zscore, weight for age	-0.91±0.94	-1.56±0.97	-1.28±0.90
Energy, kcal/d	1204±295	847±179	1110±262
kJ/d	5003±1237	3540±748	4640±1095
Total Protein, g	35.8±10.2	23.1±5.5	
33.1±7.6			
Digestibility, %	95.1±0.8	90.5±1.5	95.4±1.1
Amino Acid Score, %	94.8±5.8	85.9±7.8	87.6±8.9
Utilizable protein, g/kg/d	2.94±0.86	1.77±0.49	2.63±0.69

Apparent prevalence of inadequate intakes, %
using actual age, actual weight
using height age, actual weight
using actual age, actual weight, actual growth rate
If total food intake was in accord with fAO/WHO/UNU energy requirement **
* at mid-point of data collection (allowed age range for individual was 18 to 30 m
** assessment of adequacy of the protein:energy ratio
Based on Beaton, Calloway and Murphy (1992)

FIGURE 6
Schematic portrayal of the validation data sets generated by Willett and colleagues. Also portrayed in this schematic is the correlated error arising from the subject:instrument interface as hypothesized by Beaton.

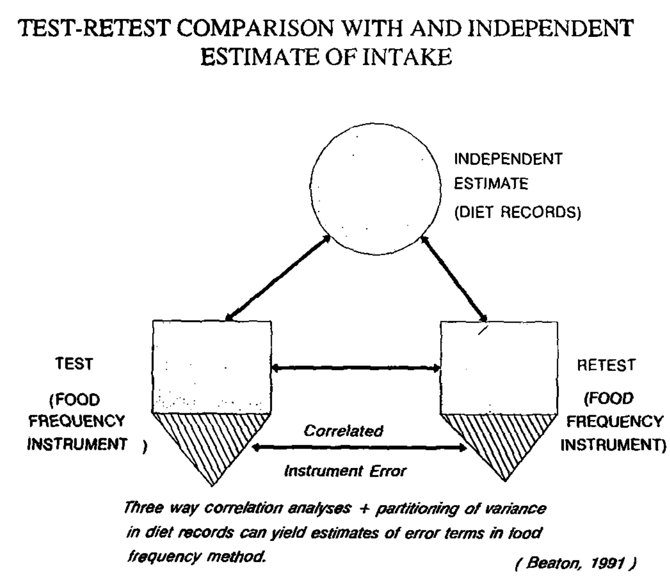


FIGURE 7

Some important features of the probability approach to nutritional assessment of estimated dietary intake.

SOME ADVANTAGES OF THE PROBABILITY APPROACH TO NUTRITIONAL ASSESSMENT OF OBSERVED INTAKE

- The probability approach has been examined and accepted on theoretical grounds.
- It opens the doors to asking new kinds of questions in the evaluation of dietary intake.
- FAO/WHO committees have begun offering the parameters that are needed for implementation. Specifically, the estimate of average requirement is the most important element of the distribution of requirements.
- Of major importance has been the presentation of requirement estimates for different levels of nutriture - iron needed to prevent anaemia - to meet functional needs - to maintain normative stores of body iron. This offers the opportunity to develop a profile of the dietary situation (prevalence of inadequacy for different levels of nutriture)
- The new technology makes possible a much broader approach to evaluation of dietary intakes and a much improved process of informing policy decisions.

REFERENCES

1. Tarasuk, V., Beaton, G.H. (1991) The nature and individuality of within subject variation in energy intake. Am J Clin Nutr 54: 464-70.
2. Tarasuk, V., Beaton, G.H. (1992) Day-to-day variation in energy and nutrient intake: Evidence of individuality in eating behaviour. Appetite 18: 43-54.
3. Beaton, G.H. (1991) Interpretation of results from diet history studies. in Kohlmeier, L. (ed) «The Diet History Method: Proceedings of the 2nd Berlin Meeting on Nutritional Epidemiology», Smith-Gordon, Nishimura, London, 1991, pp 15-38.
4. Liu, K. (1989) Consideration of and compensation for intra-individual variability in nutrient intakes in Kohlmeier, L. and Helsing, E. (eds) «Epidemiology, Nutrition and Health: Proceedings of the 1st Berlin Meeting on Nutritional Epidemiology», Smith-Gordon, Nishimura, London, 1989, pp 87-98.
5. Beaton, G.H. (1994a) Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology. Am J Clin Nutr 59(supp): 253S-261S.
6. Beaton, G.H., Milner, J., Corey, P. et al (1979) Sources of variance in 24 hour dietary recall data: implications for nutrition study design and interpretation. Am J Clin Nutr 32: 2546-59.
7. Black, A.E., Goldberg, G.R., Jebb, S.A., Livingstone, M.B.E., Cole, T.J., Prentice, A.M. (1991) Critical evaluation of energy intake data using fundamental principles of energy physiology.: 2. Evaluating the results of published surveys. European J Clin Nutr 48: 583-599.
8. Goldberg, G.R., Black, A.E., Jebb, S.A., Cole, T.J., Murgatroyd, W.A., Coward, W.A., Prentice, A.M. (1991) Critical evaluation of energy intake data using fundamental principles of energy physiology.: 1. Derivation of cut-off limits to identify under-recording. European J Clin Nutr 45: 569-581.
9. Mertz, W. Tsui, J.C., Judd, J.T., Reiser, S., Hallfrisch, J., Morris, E.R., Steele, P.C., Lashley, E. (1991) What are people really eating? The relation between energy intake derived from estimated dietary records and intake determined to be necessary to maintain body weight. Am J Clin Nutr 54: 291-5.
10. Thompson, F.E., Byers, T., Kohlmeier, L. Dietary Assessment Resource Manual. J Nutr 124 (Supp): 2245S-2317S.
11. FAO/WHO/UNU (1985) Energy and Protein Requirements. Report of a Joint FAO/WHO/UNU Expert Consultation. WHO Tech Rept Ser No. 724. Geneva.
12. NRC (1986) Nutrient Adequacy: Assessment Using Food Consumption Surveys. Nat Acad Press, Washington.
13. Beaton, G.H. (1994b) Criteria of an Adequate Diet in M.E. Shils, J.A. Olson and M. Shike (eds) Modern Nutrition in Health and Disease, Eight Edition. Vol II, Chapter 83 1491-1505. Lea and Febiger, Malvern, Pennsylvania.
14. Beaton, G.H., Calloway, D.H., Murphy, S.P. (1992) Estimated protein intakes of toddlers: predicted prevalence of inadequate intakes in village populations in Egypt, Kenya and Mexico. Am J Clin Nutr 55: 902-11.
15. Calloway, D.H., Murphy, S.P., Beaton, G.H., Lein, D. (1993) Estimated vitamin intakes of toddlers: predicted prevalence of inadequacy in village populations in Egypt, Kenya and Mexico. Am J Clin Nutr 58: 376-84.
16. Murphy, S.P., Calloway, D.H., Beaton, G.H. (1994) Schoolers have similar predicted prevalence of inadequate intakes as toddlers in village populations in Egypt, Kenya and Mexico. European j Clin Nutr.
17. Buzzard, I., Sievert, Y.A. Research priorities and recommendations for dietary assessment methodology. Am J Clin Nutr 59 (Suppl): 275S-280S.
18. Nusser, S.M., Carriquiry, A.L., Dodd, K.W., Fuller, W.A. (1995) A semiparametric transformation approach to estimating usual dietary intake distributions. J Am Stat Assoc in press.