Evaluation of Nutrient Intake: New Statistical Approaches

G. H. BEATON
Department of Nutrition, School of Hygiene
University of Toronto, Canada

SUMMARY

An approach to the application of probability statistics to the interpretation of individual dietary data or to data for populations of individuals is outlined. It is emphasized that assessment of the adequacy of nutrient intake should be based upon a judgement of the probability or risk of deficiency rather than on an "adequate or inadequate" basis as has so often happenned in the past.

By extending these concepts, guidelines for the prediction of suitable dietary protein: calorie ratios for food planing purposes have been developed. These are compared with summary data relating to existing food supplies and it is concluded that in most situations, factors limiting the aetiology of protein-calorie malnutrition than the supply of protein itself.

INTRODUCTION

It is the objective of this paper to briefly review some new thoughts on an old problem, the proper interpretation of dietary data. The concepts involved are not new but their implications have not been fully appreciated in the past.

Before proceeding, two definitions must be clarified. First, *nutritional status* is taken to mean "the health condition of an individual as influenced by his intake and utilization of nutrient, determined from the correlation of information obtained from physical, biochemical, clinical and dietary studies" (1). Clearly, nutritional status cannot be judged from dietary data alone.

Recibido: 3-1-1975.

Conversely, in the assessment of nutritional status it is necessary to make a jugement on the likelihood that the observed dietary intake is or is not adequate. In the remarks that follow, dietary *deficiency* is defined as the "situation in which the observed dietary intake is below the individual's true requirement".

For the purpose of this paper it is taken that the objective of a dietary study is the determination of whether or not deficiency exists in the individual, or in the case of a population, the proportion of the population that is deficient. It will be demonstrated that in each case, this is a matter of statistical probabilities. It might be better to think of them as the *risk to the individual* and *risk to the population.*

The Meaning of Dietary Standards

The obvious point of comparison for the interpretation of nutrient intake data is the recommended intake of the nutrient as pubished in a dietary standard. The problem is how should this be applied to the interpretation of intake data (2-7).

A series of FAO/WHO Committees have repeatedly defined the recommended intake (or more recently, th "safe level of intake") as "the amount considered sufficient for the maintenance of health in nearly all people (individuals)" (2). Similar definitions are to be found in other recent dietary recommendations (8, 9)[1]. A corollary of this is the published figure must exceed the actual requirement of most individuals. Obviously then, it is not possible to judge whether or not an observed intake is deficient by simple comparison with the recommended intake.

Individuals, even though similar with respect to age, sex, activity, body size, etc., still vary in many biological respects, including nutrient requirements. This is the well known individual variability and is illustrated in Fig. 1. When a dietary standard committee is asked to set a recommended intake, the approach has been to select an intake that meets or exceeds the needs of almost all of the individuals. If requirements are believed to be normally distributed, the practice has been to set

1. This does not hold for the description of energy requirements. Here the published figures are estimates of average needs (3, 5, 7, 8, 9).

the recommended intake at the mean + 2 standard deviations which then covers all but 2.5% of the population. This is illustrated in the figure.

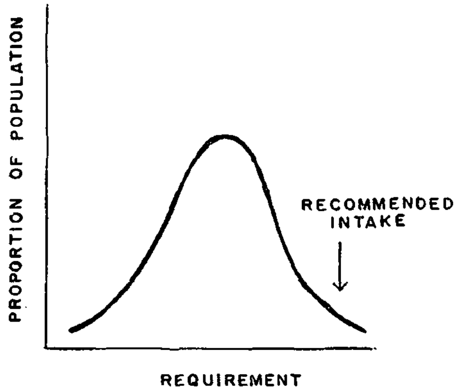

Figure 1. **Individual variability in nutrient requirement. This plot assumes that individual requirements are normally distributed about the mean requirement. The Recommended Intake has been set at two standard deviations above the mean, a level which should meet or exceed the requirements of all but 2.5% of the individuals in the population.**

The committee is really saying that at this level of intake the risk of deficiency is very low - a suitable approach for for counselling the individual.

Fig 1 demonstrates that the observation of an intake below the recommended does not necessarily mean the individual is deficient. His intake may still be above his own requirement.

The data in the graph can be transformed to a cumulative distribution describing the proportion of people with requirements above a certain level of intake (Fig. 2). But this is also a probability curve describing the risk of deficiency associated with particular levels of intake. As intake increases, the risk of deficiency decreases.

Nutrient intake data should be judged in the connotation of the probability or risk of deficiency. This, if you wish, is the crux of the "new approach".

Assessment of Nutrient Intakes

To exemplify the application of this approach, data on iron intakes of some Canadian women are shown in Fig. 3 (5). If the FAO/WHO recommended iron intake of 14 mg/day is applied to these data, it is observed that 74% of the women have intakes below the recommended level. That really doesn't tell us very much about the probable health status of these women. However, if one considers the risk, or probability of deficiency, associated with each interval of intake, it can be etimated that about 12-13% of these women are actually deficient (5). This estimate is consistent wtih biochemical and haematologic studies of similar populations (10). It is the estimate of 12% deficiency —the population risk— that is of interest in appraising nutritional status of these women, not the estimate of 74% having intakes below the recommended intake.

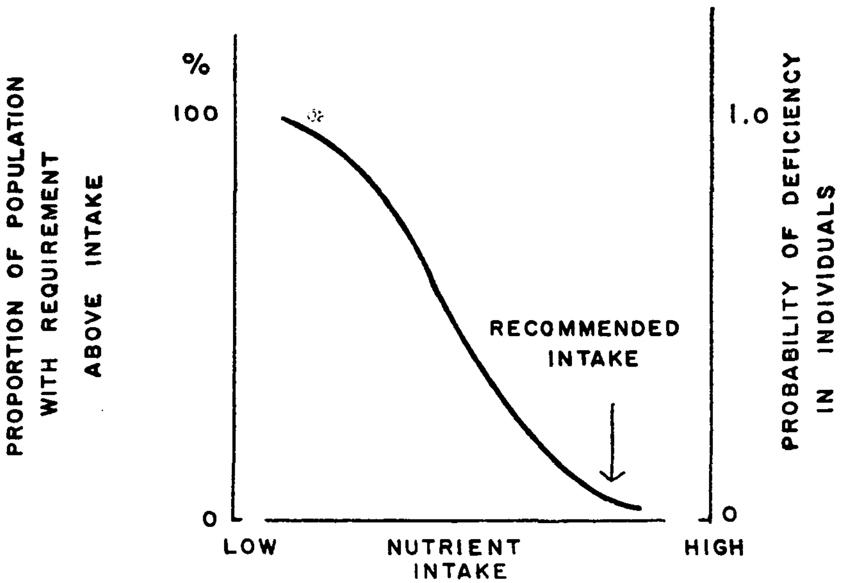

Figure 2. Probability of deficiency associated with particular intakes. The information from Fig. 1 has been replotted as a cumulative distribution showing the proportion of the population having actual requirements above a particular level. This also describes the probability of deficiency in the individual ingesting that level of intake.

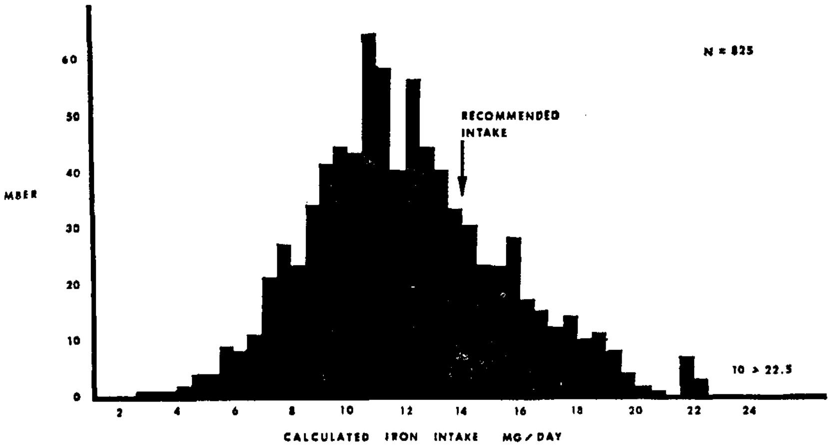

Figure 3. Distribution of iron intakes observed in groups of Canadian women (5). Individual records were collected over 37 day periods. The FAO/WHO recommended intake appropriate to these women is shown by the arrow.

A necessary prerequisite to this approach to the interpretation of dietary data is a knowledge of the nature of the distribution of individual requirements. Without this, or without some assumption about it, the risk factors associated with particular levels of intake cannot be estimated. Fortunately for our purpose, such descriptions have been published for several nutrients including: thiamine, riboflavin, niacin, iron and protein. We can begin to apply the newer approaches to these nutrients.

To employ this approach in a nutrition survey we must have a description of the average intake of the nutrient and of the variability of the intake. However there is a very particular requirement that has been emphasized by Hegsted (6). The intake data must describe the *usual* intake. The period of observation mut be long enought to eliminate the effects of day-by-day variations in intake by the individual. This effect is illutrated in Fig. 4 taken from Hegsted (6). As the period of observation increases, the apparent variability of intakes decreases. The estimation of the true variability of usual intake improves. Under Canadian conditions, three or four days of obervation is sufficient to give a constant estimate of variability of intake for many nutrients. However, in the case of nutrients, such as vitamin A, where the individual's intake may vary considerably from day to day, a much longer period of observation is required before a reliable estimate of usual intake is obtained. This must be tested in the setting of the study.

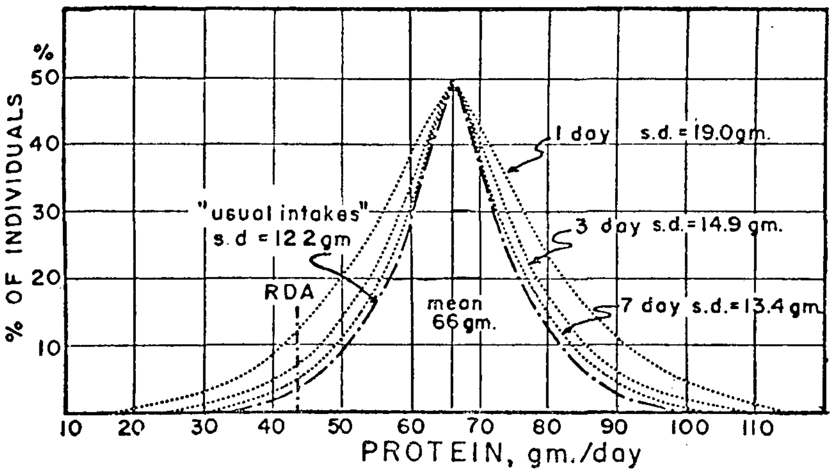

Figure 4. Effect of period of observation upon apparent variability of intake. The series of curves represent data for the same subjects observed for successively longer periods. Taken from Hegsted (6).

The statistical treatment of dietary data, involving the prediction of the risk or probability of deficiency, requires an estimate of the correlation coeficient between the individual's intake and his requirement. If the actual correlation coeficient is low, the assumption of a coefficient of 0 produces little error; if the correlation is high, a substantial error can be introduced in the statistical prediction if correlation is ignored. By taking a few precautions in the method of expressing and analysing the data to avoid spurious correlations, the conditions specified above can be met and an assumption of a very low correlation coefficient between intake and requirement can be made. For example, if thiamine intake and requirement are under consideration, it is apparent that total thiamine intake, mg/day, is likely to relate to the total amount of food consumed; it also apparent that total food consumption is one of the determinants of thiamine requirement. Thus one would expect a correlation between thiamine intake and thiamine requirement when both are expressed as mg/day. However, if both intake and requirement are expressed as mg thiamine/ 1000 kcal, the common variable of food intake is eliminated and a spurious correlation is avoided. In the case of protein, both protein intake and protein requirement would be expected to vary with body size; if both are expressed as g protein/ kg body weight, the common variable is eliminated and again a spurious correlation is avoided. Under these conditions, the correlation between intake and requirement apprears to be very low (7).

For similar reasons, this approach should be applied to relatively homogenous subgroups of the population —young children, menstruating women, adult men, etc.— rather than to the total population.

The statistical techniques themselves involve the application of the bivariate distribution of intake and requirement to predict the incidence of deficiency among the members of the population. This has been described elsewhere (4, 7, 11).

One might ask why this approach has not been applied in field studies. The answer is probably three-fold. Until relatively recently, estimates of the variability of nutrient requirements were not available; even to-day they are available for only a few nutrients. Without such estimates the approach is impossible. Second, over the years, because of the apparent discrepancy between conclusions drawn from dietary data and from biochemical and clinical data, there has grown up an inherent distrust of dietary studies. The conclusions can now be drawn on a more appropriate basis, but the distrusts will take some time to overcome. Finally, this approach to the analysis of dietary data places special requirements on the type of data needed. Specifically, 1 day data and household data are of very limited value; these are the more common types of data now collected.

Again I suggest that we are now in a position to greatly improve the interpretation of dietary data, and hence to improve the assessment of nutritional status. However, we must be prepared to modify the design of nutrition surveys. It is essential that future studies begin to focus on the *usual* intake of the *individual,* not the average intake of a household.

The Special Case of Protein

Unfortunately, the particular approaches described above have limited applicability to protein. Protein deficiency, at the tissue level can develop as the result of an inadequacy of either protein intake or caloric intake or some combination of the two. Our knowledge of the quantitative aspects of these interactions is too meagre to permit meaningful predictions of the incidence of deficiency from dietary data (7).

However, Beaton and Swiss (7) have recently applied the principles outlined in this paper in a different way. Rather than attempting to predict the prevalence of deficiency in existing populations, attention was directed to a planning problem.

The objectives of food and nutrition planning are:

1. to provide a sufficient amount of food to meet energy needs,
2. to ensure that this food is of adequate quality to meet nutrient requirements when energy needs are satiated.

The relevant question then becomes, what concentration of dietary protein, as % calories, would meet the protein requirements of almost all individuals in the population when they ingested enough food to meet their particular energy needs? This problem was solvable on the basis of existing knowledge and the assumptions inherent in the FAO/WHO Report on Energy and Protein Requirements (3).

The data shown in Table 1 were derived from this study. The displayed figures were derived for preschool children, children and adolescents; adult figures were slightly higher[2]. These estimates appear to have wide empiric applicability; seemingly they are only marginally affected by growth rate[2] and are not affected by such parameters as body size.

Across the top of the Table are two population risk categories, a predicted prevalence of either 1% or 2.5% "deficiency". The conventional recommended intake is based upon a 2.5% risk at the level of the individual.

The two variables of concern in the population, the variability of the protein concentration among the usual intakes of individuals in the population, and the quality of the dietary protein are displayed in the Table.

2. No attempt has been made to examine this relationship for pregnant or lactating women of for young infants.

TABLE 1
CRITERIA OF ADEQUACY OF AVERAGE PROTEIN:
CALORIE RATIO IN POPULATION DATA

Coefficient of Variation of Dietary Protein Concentration %	To Meet Needs of All but 1% Individuals	To Meet Needs of All but 2.5% of Individuals
	Protein with Utilization = 80% that of Egg or Milk	
15	7.9[2]	7.2[2]
20	9.2	8.1
25	11.4	9.4
	Protein with Utilization = 70% that of Egg or Milk	
15	9.1	8.3
20	10.5	9.3
25	12.1	10.7
	Protein with Utilization = 60% that of Egg or Milk	
15	10.6	9.6
20	12.3	10.8
25	15.3	12.6

1. Refers to the relative biological utilization of the dietary protein (3).

   This might be calculated as $\frac{\text{NPU (diet)}}{\text{NPU (milk or egg)}} \times 100$
   or by some other biological assay in which both the diet and egg or milk were assayed under identical conditions. An approximation of the value may be obtained by use of amino acid scores as set out in the FAO/WHO report (3).
2. Protein calories as % of total calories, assuming a value of 4.0 kcal/g protein.

The presentation in the Table should prove useful in answering some very practical questions. Consider the situation of a country, region or sub-population in which protein caloric malnutrition is known to exist. In all probability, the total level of food consumption in this population, or among the members of the high risk sub-population, will be low. The planner can anticipate that more food will be needed if the problem is to be rectified. However, should he plan on more of the same type of food? Could he provide sources of energy (sugar or fat) that might be cheaper than increasing the general food supply? Or must the provide an increased concentration of protein in the diet? Table 1 should offer a reasonable answer to the question, at least, the best answer we can now offer. If the concentration of protein in the existing diets approximates or exceeds the appropriate values in the Table, then increasing the types of foods now available would meet energy and protein needs. This appears to be the case in many if not most areas for which data are available (7). However, if the concentration of protein is below the values given in the Table, as it would be in a cassava-eating area, then protein rich foods must be added to raise the concentration. The Table can also be used to assess the danger of adding sugar and fat which would, in effect, dilute the protein concentration. It is a feasible approach only if the initial concentration is appreciably higher than the value shown in the Table.

Comment on Implications in Nutritional Studies

In areas or population groups in which the average or *per capita* protein: calorie concentration is above the critical levels suggested in Table 1, PCM may still exist. However, it is suggested that the prime cause is unlikely to be inadequate types of foods in terms of their protein concentration. Rather, the occurrence of PCM is more likely to be the result of inadequate amounts of food, or the result of particular practices which keep the food away from the affected children.

It must be recognized that this approach does *not* predict the prevalence or risk of protein deficiency in the existing population. Indeed, at the present time, it is the contention of this author that appropriate statistical techniques to do that do not exist.

Rather, the present approach simply establishes a procedure for making qualitative judgements about the food supply itself, a judgement which may help to separate out those few areas of the world where protein fortification or amino acid fortification would seem to have a role to play. Typical data on observed intakes, derived from dietary surveys, are shown in Table 2; these may be compared with the figures given in Table 1.

In most areas of the world, it would appear that the basic problem remains two-fold: (a) a need to increase the *amount of food* effectively reaching the individual consumer, and (b) a need to ensure that there are appropriate and suitable foods *sufficiently high in caloric and nutrient density* for the feeding of young infants and children.

TABLE 2

APPARENT AVERAGE PROTEIN CONCENTRATIONS OBSERVED IN DIETARY SURVEYS CONDUCTED BY THE ICNND

Area and group	Average protein concentration, % kcal
Latin America	
Bolivia,[1] families	12,2
Chile, families	12.6
Colombia, families	10.2
Ecuador[1], families	13.0
Northeast Brazil[1],	
Families	14.0
Pregnant wmen	16.5
Infants under 2 years	11.9
Uruguay[1]	
Families	14.0
Children, 3-4 years	21.1
Children, 1-2 years	21.1
Caribbean	
Trinidad, families	13.2
St. Lucia, families	12.8
St. Kitts, families	17.1
Nevis, families	17.3
Anguilla, families	13.6
Alaska, males, all ages	29.3
Middle East	
Ethiopia, families	12.4
Jordan,[2] families	12.6
Lebanon,[2] families	12.6
Far East	
Burma, families	9.1
East Pakistan, families,	
Rural	10.2
Urban	11.2
Malaya	
Families	11.2
Children	10.7
Thailand, families	10.7
Vietnam, families	
Vietnamese[1]	10.4
Highlanders	10.4

[1]. Two or more regions studied; concentrations generally comparable.

[2]. Refugees and non-refugees studied; concentrations generally comparable.

The protein: calorie ratio data presented in this paper is a special application of the concepts described in the earlier part of the paper. In itself it is *not* a technique to be used in the appraisal of nutritional status. An example of the approach to be used in the assessment of nutritional status was given with reference to iron.

The major point that emerges from these examples is that dietary data, whether derived for the individual or the population —like biochemical, anthropometric and clinical data— must be interpreted in terms of the *probability* that inadequacy does or does not exist.

The few examples provided in this paper reflect some new approaches to the analysis of dietary data. These approaches are admittedly imperfect; they involve assumptions that are not yet well tested. They call for types of data that are not now being collected. Nevertheless, it seems clear that they provide interpretations that are much more meaningful and much more useful than do existing approaches. It is to be hoped that of this type will be applied in future studies.

Ultimately, if the intent is to truly judge nutritional status, it will be necessary to use statistical approaches that begin to combine the parameters of nutritional status at the level of the individuals within the population, not at the level of the population as a whole.

RESUMEN

Evaluación de la ingesta de nutrientes: Nuevos enfoques estadísticos

Se expone un enfoque relativo a la aplicación de la estadística de probabilidades para la interpretación de la información dietética individual o de la dimanante de grupos de población. Se enfatiza que la adecuación de la ingesta de nutrientes, debe ser apreciada a la luz de la probabilidad o riesgo de deficiencia en vez de sobre una base de "adecuado" o "no adecuado", como tan frecuentemente se ha hecho hasta ahora.

Extrapolando estos criterios se han desarrollado orientaciones para la adecuada predicción de la relación proteína/caloría, a ser usada en planificación alimentaria. Una comparación entre este concepto y el suministro existente de alimentos, permite concluir que en la etiología de la malnutrición proteínico-calórica, muchas veces los factores que limitan la disponibilidad y el consumo de los alimentos, pueden ser más importantes que el suministro mismo de proteínas.

BIBLIOGRAPHY

1. Christakis, G. Nutritional Assessment in Health Programs. Am. J. Pub. Health 63, Nov. 1973. Supplement.
2. FAO/WHO Joint Expert Committee on Nutrition. Eigth Report: Food Fortification, Protein-Caloric Malnutrition. Wld. Hlth. Org. Techn. Rep. Ser. No. 477, Geneva; FAO Nutr. Meeting Rep. Ser. No. 49, Rome. 1971.
3. Joint FAO/WHO Ad Hoc Expert Committee. Report on Energy and Protein Requirements. Wld. Hth. Org. Techn. Rep. Ser. No. 522, Geneva; FAO Nutr. Meetings Rep. Ser. No. 52, Rome. 1973.
4. Lörstad, M. H. Recommended Intake and its Relation to Nutrient Deficiency. FAO Nutr. Newsletter 9: No. 1. p. 18. 1971.
5. Beaton, G. H. The Use of Nutritional Requirements and Allowances. In. Proc. West Hemisphere Nutr. Cong. III, P. L. White (ed.) Futura Press. p. 356. 1972.
6. Hegsted, D. M. Problems in the use and interpretation of the recommended dietary allowances. Ecol. Food Nutr. 1: 255 1972.
7. Beaton, G. H. and Swiss, L. D. Evaluation of the nutritional quality of food supplies: production of "dsirable" or "safe" protein-calorie ratios. Am. J. Clin. Nutr. 27: 485. 1974.
8. U. S. Recommended Dietary Allowances, 8th Edition. Food and Nutrition Board, National Research Council. National Academy of Sciences, Washington. 1974.
9. Dept. of Heath and Social Security. **Recommended intakes of Nutrients for the United Kingdom**. Rept. on Public Health and Med. Subj. No. 120. Her Majesty's Stationery Office. 1969.
10. Beaton, G. H. Epidemiology of Iron Deficiency in **Iron in Biochemistry and Medicine**. A. Jacobs and M. Worwood (ed.) Academic Press, London. 1974.
11. Lörstad, M. H. On Estimating Incidence of Undernutrition. FAO Nutrition Newsletter. 12, No. 1, p. 1. 1974.