REFERENCES

Accessory Food Factors Committee (1945). *War Memor. med. Res. Coun., Lond.*, no. 14.
Department of Health (Republic of Ireland) (1949a). *Methods of Dietary Survey and Results from Dublin Investigation.* Dublin: The Stationery Office.
Department of Health (Republic of Ireland) (1949b). *Dietary Survey of the Congested Districts.* Dublin: The Stationery Office.
Department of Health (Republic of Ireland) (1950a). *Dietary Survey of Large and Small Towns.* Dublin: The Stationery Office.
Department of Health (Republic of Ireland) (1950b). *Dietary Survey of Farming Families.* Dublin: The Stationery Office.
Geary, R. C. (1950). *Brit. J. Nutrit.* **4**, 274.
Magee, H. E. (1944). *Mon. Bull. Minist. Hlth*, **3**, 146.
Magee, H. E. (1945). *J. med. Ass., Eire*, **17**, 98.
Sydenstricker, V. P. (1944). *J. Amer. diet. Ass.* **20**, 4.

Statistical Plans and Methods of the Irish National Nutrition Survey

By R. C. GEARY, *Director, Central Statistics Office, Dublin*

I propose dealing almost exclusively with some methodological aspects of the National Nutrition Survey from the narrowly statistical point of view, leaving statistical inferences and general comments on the results to other speakers. The Central Statistics Office was privileged to collaborate with the Department of Health, but not in so intimate a way as to render it unbecoming for us to congratulate the Minister, Dr Deeny and Dr Hourihane warmly on the various excellences of the inquiry or for the Office to address itself frankly to a criticism of the methods used, even though we have a measure of responsibility for these methods and the wisdom is after the event. What emerges in this short paper is commentary rather than criticism. This is a first inquiry and one of the most useful objects of a first inquiry was indicated by the late Lord Stamp to a colleague of mine when he remarked (in regard to a statistical inquiry in which Ireland was a pioneer): 'While it is interesting to know just what you have done, it would be still more enlightening to hear what you would not do if you were repeating the inquiry.' It is well to remember that considerable advances have been made in the sampling techniques applicable to surveys of this kind since the present survey was planned.

The Central Statistics Office helped in the selection of the families to be surveyed, in making the computations from the machine tabulations and in preparing the tables in the report for printing.

Selection of families

It was decided that the survey should extend to 2500–3000 families, i.e. to about one in 250 of the private families in the country. The number of families surveyed in each of six main categories is shown on p. 275. The Dublin City sample was selected from four wards, Mountjoy, Rotunda, Crumlin and Drumcondra. The families in each ward were selected on a random basis from the 1943 Register of Population. Assignment to the three grades, slum, artisan and middle-class, was based on the characteristics of

individual families, not on areas of residence, though the majority of the slum families resided in Mountjoy and Rotunda, the artisan families in Crumlin and the middle-class in Drumcondra.

Type of family	Number of families in sample
Families living in Dublin City	500
Families living in thirteen large towns	329
Families living in forty-nine small towns	516
Farming families	948
Families of farm labourers	177
Families living in congested districts	200
Total	2670

It should be pointed out at this stage that, in general principle, the Department of Health decided not to adhere rigidly to the tenets of random sampling in the town areas but to bias the sample generally towards the poorer sections of the community. Thus, as will presently appear, the average social grade of the Dublin sample was lower than the general average for the city. In the small and large towns, families larger than the average were selected. This bias is also evident in the large number of towns included in the sample from the poorer western districts. In the rural division of the inquiry the same intention is evident in the special investigation of the congested districts. While the statistician may cavil at this bias he cannot withhold sympathy from its intention. The object of official inquiries into social amenities is not so much to ascertain unbiased national averages of the various factors as to identify the sectors where remedial action is necessary and to measure accurately the deficiencies, if any. These inquiries are expensive and there must seem to be but little point in securing proportional representation in the sample of the more affluent sections of the community in regard to which one knows in advance no problem is likely to arise. Furthermore, using census data on, for example, geographical distribution of the population, social status and size of family, as controls, it is possible to correct the bias of the estimates in a considerable degree: some examples will presently be given. Averages derived from more or less homogeneous classes, social, geographical, family size or income, are of more importance as being more revealing than general averages.

The thirteen large towns include all towns (except Dublin) with populations exceeding 10,000 and two towns with population just under 10,000. In the large towns the sample was roughly proportional to the population of the town but eight families were selected from each small town, whatever its population. Furthermore, a substantial transfer of families (deemed in the first instance to be 'farming families') to 'large' and 'small towns' occurred as the field work progressed when it was ascertained that such families obtained a high proportion of their income from non-agricultural occupations.

In town areas, arrangements had to be made for substitution in case the original family scheduled for investigation was inaccessible or non-co-operative. An elaborate series of instructions for finding substitutes was given to each investigator. The provision of substitutes was another reason for an increase in the number of families sampled in towns. Should an investigator make a mid-week call on a family who had been provided

with a log-book and obtain no reply, he would immediately find a substitute; sometimes it happened that he would obtain log-books from both the original family and the substitute. The extent of substitution was about 25 % in town areas and it is not likely that such a degree of substitution imparted any appreciable bias to the results. It may be surmised that any such bias was well within the random sampling error.

Farming families were selected from the Agricultural Statistics lists. A number of District Electoral Divisions (D.E.D.) were chosen in each county roughly proportional to the county's farming population and within each D.E.D. ten names (with ten alternatives) were supplied to investigators. The selection was made in each county in such a manner that the sample numbers in each size of farm group were proportional to those of the whole county. For the convenience of investigators, the D.E.D.'s were selected adjacent to small towns. It is unlikely that appreciable bias was introduced by this departure from the principle of randomness.

Farm labourers' families were investigated concomitantly with the farming families. In each county the number selected was proportional to the number of married male agricultural labourers. Selection was made by obtaining the addresses of the nearest farm labourers at the first farms visited until the required number in the county was attained.

A special inquiry was extended to the congested districts where the density of farming population is high and where conditions of life are hard. Twenty D.E.D.'s were selected in counties Kerry, Galway, Mayo, Leitrim and Donegal, the number in each county being proportional to the county's share in the D.E.D.'s deemed to be 'congested'. These counties were surveyed independently in the sample of farming families. The sample D.E.D.'s were selected chiefly with reference to their accessibility for investigators but, as far as possible, only the poorer type of congested district was included. The names and addresses (with ten substitutes) were selected in each D.E.D.

To sum up as regards the selection of families, in town areas a deliberate bias towards poorer and larger families was introduced into the sample, but the sample may be considered random, and therefore representative, for farming and farm labourers' families. The congested districts sample was derived from the poorer parts of these districts.

Comparison of the uncorrected and corrected sample averages

Since there was a bias towards the poorer districts in Dublin City, it may be anticipated that the true average consumption of foods in Dublin City as a whole is higher than the figures revealed by the sample. It will be shown, however, that the bias introduced into the sample is small and, indeed, well within the sampling error.

In order to correct the sample figures for bias an analysis was made of the occupied males in each of the three family types (slum, artisan and middle-class). Percentages in the sample and in the city as a whole (Census of 1936) in four main occupational groups are shown on p. 277.

The occupational distribution of the occupied males in the three family divisions of the sample was compared with the occupational distribution of all occupied males in Dublin, according to the Census of 1936, and multipliers 3, 7 and 5 were estimated, by

least square procedure, as weights to be applied to the averages of slum, artisan and middle-class divisions, respectively, in order to find averages for the city of Dublin as a whole.

	Sample %	Census %
Professional, clerical and commercial	25·6	31·4
Skilled manual workers and men in supervisory positions	22·7	25·2
Semi-skilled manual workers	26·7	22·3
Unskilled manual workers	25·0	21·1
Total	100	100

Table 1 shows a comparison of the sample averages with those found for Dublin City as a whole by applying the weights to the three divisions of the sample.

Table 1. *Average daily consumption of foods and intake of calories per diet-head in Dublin City*

Food	Sample average	Weighted Dublin average
Milk (pt.)	4·8	5·1
Cheese (oz.)	1·2	1·2
Eggs (no.)	4·0	4·3
Fats (oz.)	9·6	9·7
Meat (oz.)	33·3	34·4
Fish (oz.)	3·9	3·9
Flour and bread (oz.)	75·8	74·4
Other cereals (oz.)	8·4	9·7
Potatoes (oz.)	82·7	82·1
Vegetables (oz.)	27·4	30·0
Fruit and fruit products (oz.)	9·3	11·5
Sugar (oz.)	12·8	12·9
Preserves (oz.)	3·9	4·3
Miscellaneous (oz.)	0·6	0·6
Calories	2641	2685

It will be observed that the unbiased averages for Dublin City are not markedly different in most cases from the sample averages. The greatest percentage difference in consumption occurs with fruit and fruit products where the difference is about 20 %. A further investigation (discussed later) shows that the differences are well within the sampling errors and that, therefore, the sample averages may be taken as representative of the average consumption in the total population of Dublin.

The families sampled in the towns other than Dublin were not subdivided by social status or type of area of residence and apart from the omission of professional families there was no bias on this account. There was, however, a marked bias towards the larger families. It was possible by using as weights the proportions of persons in each family group in the town areas (excluding Dublin), as found from the Census of Population, to eliminate the effect of the bias and present corrected averages for large and small towns. A comparison between the uncorrected averages and the corrected averages is shown in Table 2.

Though the uncorrected averages are in all instances somewhat lower than the corrected averages, there is very little difference in the figures for flour and bread, potatoes, meat, sugar and cheese, but correction has the effect of increasing the intake of calories figure by 6 % for each. The foregoing are but a few samples of how the

Table 2. *Uncorrected averages and averages corrected for family size in large and small towns*

Food	Large towns average		Small towns average	
	Uncorrected	Corrected	Uncorrected	Corrected
Milk (pt.)	4·4	4·9	5·1	5·7
Cheese (oz.)	1·2	1·2	0·8	0·8
Eggs (no.)	3·4	3·7	3·7	4·3
Fats (oz.)	8·9	10·0	8·7	9·1
Meat (oz.)	31·2	32·7	25·1	28·5
Fish (oz.)	2·5	2·9	3·1	3·4
Flour and bread (oz.)	82·1	84·9	79·4	80·6
Other cereals (oz.)	8·1	9·4	6·7	7·0
Potatoes (oz.)	102·7	106·5	110·0	112·0
Vegetables (oz.)	25·6	28·3	22·6	24·9
Fruit and fruit products (oz.)	6·7	7·3	5·7	6·9
Sugar (oz.)	9·0	9·4	11·6	11·9
Preserves (oz.)	4·0	4·6	4·4	5·0
Miscellaneous (oz.)	0·7	0·8	0·4	0·5
Calories	2637	2804	2601	2734

downward bias in the samples for towns may be corrected using control data. The full corrections are given in the Appendix to Part 1 of the National Nutrition Survey (Department of Health (Republic of Ireland), 1949a).

A point of theory as to cost of inquiry

Inquiries by the interview method are expensive and the present inquiry was no exception. There is an aspect of sampling theory appropriate to these inquiries which tends to be overlooked but which has an important bearing on the cost. In its simplest form it can be expressed in these terms: even if each individual return is absolutely accurate, an estimate based on a mean of a random sample of n individuals is subject to sampling error, namely $\sigma_x/\sqrt{n}$, where σ_x is the standard deviation of the quantity x in the population and the population is indefinitely large. The cost of the inquiry depends on several factors, of which it is necessary to signalize two for the present purpose, namely (1) the number in the sample and (2) the degree of accuracy deemed essential in each return included in the sample. Unfortunately, in regard to (2), we do not know how cost per return varies with the individual accuracy standard adopted. At a guess one might suggest a hyperbolic law

$$c\sigma_e = A, \text{ a constant,} \tag{1}$$

connecting c, the cost per return, with σ_e, the individual error (e) standard deviation. The observed value $y = x + e$ so that $\sigma_y^2 = \sigma_x^2 + \sigma_e^2$ and the standard deviation of the mean of n is $\sigma_y/\sqrt{n}$. The total error variance is

$$\frac{\sigma_y^2}{n} = \frac{1}{n}(\sigma_x^2 + \sigma_e^2), \tag{2}$$

the total cost being a given constant B, so that

$$nc = B. \tag{3}$$

It can now be shown, by a particular series of examples, how the cost of inquiry

increases with increasing precision requirements in the individual return, given as constant the overall error standard deviation $\sigma_y/\sqrt{n}$. Let $A = 1$, $\sigma_x = 10$, $\sigma_y/\sqrt{n} = 1$. From the foregoing formulas the total cost $B = nc$ is as shown on the second line of the following table, corresponding to the four individual error standard deviations shown on the first row.

Individual error S.D. σ_e	15	10	5	1
Total cost nc	21·7	20	25	101

The steep increase in cost with increasing individual accuracy will be noted, showing that it is wasteful to strain after meticulous precision in individual returns. Actually the method of inquiry which entails an individual error S.D. (σ_e) of 10 is the most economical. Greater overall accuracy will be obtained by using the funds available to secure the maximum number of returns even if, in the less meticulous character of the questionnaire, less personal attention required by the interviewers, and less querying, individual accuracy is discarded as the ultimate objective. Considerable economy might be effected by using returns by mail from a fraction of the sample which might be able to cope unaided with the simplified questionnaire.

This analysis is subject to the assumptions, however, that equation (1) holds approximately and that the individual errors are unbiased. It might be possible to investigate the truth of the latter assumption (and perhaps even to throw some light on the former) by using two-phase sampling technique which in this connexion would involve meticulous assessment of a subsample of the original sample and comparison of the corrected data with the data originally returned in the subsample.

Random sampling errors of estimate

As an experiment which could be extended, the standard deviations of the estimates for four foodstuffs in the three Dublin zones have been computed. The results are shown in Table 3.

Table 3. *Mean values with their standard deviations for the weekly consumption of four foods in Dublin*

Food	Slum (188 families)	Artisan (210 families)	Middle-class (102 families)
Fats (oz.)	9·7 ± 0·32	9·5 ± 0·21	9·9 ± 0·31
Sugar (oz.)	12·6 ± 0·37	13·1 ± 0·19	12·7 ± 0·43
Eggs (no.)	3·2 ± 0·22	3·8 ± 0·21	5·6 ± 0·30
Potatoes (oz.)	84·2 ± 2·93	82·6 ± 1·91	80·2 ± 3·82

The inference is, of course, that the true means for the whole populations concerned probably (probability 19/20) lie within twice the range indicated in the table.

Relation between income and expenditure on food

In the published reports (Department of Health (Republic of Ireland), 1949a, b, 1950a, b) all the data represent averages per diet-head weekly for different categories of households. Very useful researches could be made into such relationships as these

between consumption of various foodstuffs, or between foodstuffs and income using individual returns. The 900 returns available from the Survey up to November 1947 were utilized for the determination of weights for the official Interim Cost of Living Index Number in a manner described in detail elsewhere (see *Irish Trade Journal and Statistical Bulletin*, June 1948, p. 93). The method used involved incidentally the determination of expenditure on food per head (f) for Dublin as a linear function of family income per head (i), family fuel and light expenditure (l), rent (r), dependent children (d), other persons in family (w), as follows:

$$f = 8\cdot3453 + 0\cdot2382i - 0\cdot0447l + 0\cdot0238r - 0\cdot2784d - 0\cdot2932w.$$

All values are in shillings/week. The equation shows that factors other than food affect the relationship in comparatively minor degree. As illustrating Engel's Law as applied to the non-agricultural employee class in Ireland in 1946–7 the equation shows that expenditure on food as a proportion of total income was 54% when the income was 100s./week, 44% when the income was 147s./week (the actual average) and 39% when the income was 200s./week. The equation also shows that other things being equal the larger the size of family (even though the larger the family the greater the number of earners) the lower the average expenditure on food per head.

Suggestions for future inquiries

When this inquiry was first mooted, the Central Statistics Office naturally gave some thought to the possibility of grafting to the nutrition survey a full-scale family budget inquiry (designed primarily to establish a series of weights for a new series of cost of living index numbers), but this was judged inexpedient. At the same time the experience gained at the nutrition survey should prove useful when the family budget survey is instituted. On the other hand, data of considerable value from the nutrition point of view should be obtainable from a family budget inquiry. A specific recommendation that family budget inquiries should be designed in consultation with nutrition experts was made at the Seventh International Conference of Labour Statisticians held under I.L.O. auspices last autumn. The questions in regard to food consumption will necessarily be less elaborate than those asked at the present inquiry and the results will be more subject to what has been termed above 'individual error'. It may, for instance, be necessary to ignore changes in stocks and to confine questions to purchases.

Following are some suggestions for future inquiries, based on the experience of the present first inquiry.

(1) The inquiry should be absorbed in a full family budget survey to the conduct of which the recommendations of the Seventh International Conference of Labour Statisticians would apply.

(2) Less meticulous accuracy in individual returns will permit of a considerable increase in coverage (i.e. number of returns) for given cost, with increased accuracy in overall estimates in consequence.

(3) Ascertainment of full expenditure (and not only expenditure on food, rent, fuel and light) will improve the accuracy of the family income figures, so necessary for the analysis of nutrition data.

(4) Under a less meticulous system it should be possible to obtain more than 1 week's consumption from each family surveyed so as to correct for seasonal variation; consumption should be analysed systematically by months, quarters or seasons throughout the period of inquiry, having regard to the very striking results obtained for the congested districts from the present inquiry.

(5) Random sampling standard deviations might be supplied with the main estimates.

I understand that the Department of Health is considering the publication of a concluding part to the Irish National Nutrition Survey in addition to those parts relating to the five sections of the inquiry. This part would be, no doubt, designed primarily to give a description of sampling methods used and other methodological aspects of the inquiry not dealt with in the other parts. It may also be intended to include global totals for the country as a whole and perhaps for all town areas and all rural areas separately. If this is so, it would be well to correct the figures for possible bias, in so far as this can be done, using, perhaps, the methods of which examples have been given in this contribution. It also seems desirable to give some indication of the magnitude of the random sampling errors to which these global estimates are subject.

REFERENCES

Department of Health (Republic of Ireland) (1949a). *Methods of Dietary Survey and Results from Dublin Investigation*. Dublin: The Stationery Office.
Department of Health (Republic of Ireland) (1949b). *Dietary Survey of the Congested Districts*. Dublin: The Stationery Office.
Department of Health (Republic of Ireland) (1950a). *Dietary Survey of Large and Small Towns*. Dublin: The Stationery Office.
Department of Health (Republic of Ireland) (1950b). *Dietary Survey of Farming Families*. Dublin: The Stationery Office.

The Results of the Irish National Nutrition Survey

By W. J. E. JESSOP, *Royal College of Surgeons in Ireland, St Stephen's Green, Dublin*

The results of the first four sections of the Dietary Survey phase of the Irish National Nutrition Survey have been published in some ninety-five separate tables, a number of them of considerable length. These tables are available (Department of Health (Republic of Ireland), 1949a, b, 1950a, b) but quite obviously it would be impossible to consider each of them in detail. I have decided that, in the time at my disposal, I could best present these results by attempting to answer the following questions: (1) What foods do the different sections of the population eat? (2) How is the consumption of these foods affected by variations in economic circumstances? (3) How does the intake of the several nutrients by the various classes surveyed compare with recognized standard requirements?

I think these questions will be found to be answered by Tables 1 and 2. I will discuss briefly the figures in these tables but it will not be possible in the time available to enter into comparisons between our results and those of other surveys. Comparisons of this kind are, of course, often of great value and we would be very interested to hear comments on food habits in Ireland from this point of view.