Foods and diseases

E. G. KNOX

From the Health Services Research Centre, Birmingham

SUMMARY An examination was made of the statistical correlations between the main foodstuff and nutrient intakes and the chief causes of mortality in 20 different countries, comprising 17 in Europe, and Canada, USA, and Japan. Subsidiary examinations were made of the effects of including and excluding Japan, and of the effects of various statistical standardisation procedures. Complex food patterns were identified and related both to geographical latitude and to levels of affluence; these, in turn, were related to complex patterns of mortality. Criteria for drawing special attention to specific associations were identified, based partly on statistical significance tests and also on strength-of-association yardsticks supplied by diseases with known causes. Findings suggesting causal interpretations were: (*a*) alcohol intakes and cirrhosis of the liver, cancer of the mouth, and cancer of the larynx; (*b*) total fat intakes and multiple sclerosis, cancer of the large intestine, and cancer of the breast; and (*c*) beer and cancer of the rectum.

International variations in the incidence of illness provide striking epidemiological characteristics of many diseases. Examples are: the relative rarity of anencephalus in France compared with England, the high incidence of carcinoma of the breast in North America compared with Japan, the temperate zone distribution of multiple sclerosis, the high incidence of ischaemic heart disease and of carcinoma of the large intestine in Europe compared with Africa.

Transfers of complex living patterns from one country to another—for example, from Europe to the emerging middle classes of developing African and Asian countries—have been followed by the disease patterns of the donor country, and studies of migrants have in other cases demonstrated their progressive acquisition of the disease patterns of the new domicile. Studies have implicated environmental rather than (purely) racial and genetic factors and the dietary component of the changed culture reasonably excites suspicion.

A wide range of geographical associations between specified food and nutrient intakes and disease incidences continue to be reported, but the complexities of both disease and dietary patterns, each with its own set of internal correlation, and each with a set of associations with non-dietary components of life-style, make critical interpretation a formidable task. Certainly, the arbitrary selection of any single food/disease pair, and its examination for a positive or negative association across a range of different countries or different times, can scarcely do justice to the complexity of the background from which the particular example was picked. It would seem necessary, at least, to examine a range of foodstuffs and a range of diseases besides the ones in question. This study is an investigation in this sense and is based upon the availability of two sets of international data from a common range of countries, which do not seem previously to have been associated in a systematic way.

Materials and method

The Organisation for European Co-operation and Development (OECD) produces at intervals a series of tabulated *Food Consumption Statistics* and the data used for this study are the figures for 1960-68 (Organisation for European Co-operation and Development, 1970). The tables refer to 20 countries namely: Canada, United States of America, Japan, Austria, Belgium/Luxembourg, Denmark, Finland, France, Germany (FR), Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and Yugoslavia.

A large variety of foodstuffs and foodstuff-groupings is listed, 113 altogether, but the listing is not uniform and 57 commodities are indexed as occurring in only a 'small number' of countries. Examples in this class are buckwheat, molasses, tapioca, goat's meat, cottage cheese, shell-fish, and sea weeds. An abridged list was constructed, mainly of foods not so indexed, although a few were included because of possible aetiological connections with disease. The values used were mean grams

consumption per caput each day for the years 1960-64, or, if recording started later, the earliest available figure. The foods were: wheat, rye, total bread grain, barley, oats, total coarse grain, rice, total cereals, refined sugar, honey, potatoes, pulses, vegetables, citrus fruits, fresh non-citrus fruits, canned fruit, dried fruit, nuts, fruit juices, cocoa beans, beef, veal, lamb/mutton, pig meat, horse meat, poultry, other meat, edible offals, total meat, eggs, cows' milk, skimmed milk, cream, condensed milk, dried whole milk, dried skimmed milk, cheese, fresh and frozen fish, salt/cured/dried fish, canned fish, olive oil, other vegetable oils, total vegetable oils, marine oils, lard, margarine, butter, wine, and beer. Total calories from all sources, from animal and vegetable sources separately, and within each class, from fat and protein, were also prepared. Total grams weight of the main foodstuffs (excluding wine and beer) was used as an index of total dietary intake. The latitude of the capital city of each country was also recorded.

The World Health Annual Statistics for 1970 (World Health Organisation, 1973) was used as the source of disease-mortality data. All the countries in the OECD list were matched except for the Belgium/Luxembourg economic union, for which Luxembourg was the nearest equivalent. The age specific mortality rates per 100 000 were classified in accordance with the 'A' list of the 8th revision of the *International Classification of Diseases* (World Health Organisation, 1967) for 137 non-violent and 13 violent causes of death.

Once more, an abridged list was prepared. Diseases with very low mortalities were excluded as were most of the infective and violent causes although respiratory tuberculosis, suicide, death from road accidents and from drowning, were included. The main neoplastic and circulatory disease groups were included as were goitre, thyrotoxicosis, diabetes, psychosis, neurosis, multiple sclerosis, epilepsy, chronic bronchitis, peptic ulcer, appendicitis, cirrhosis, gall bladder disease, chronic nephritis, hyperplasia of prostate, toxaemia of pregnancy, arthritis and spondylitis, spina bifida, and congenital heart disease.

In many cases the mortality within a single age group was used, most frequently the age group 55 to 64 years, within which there were usually sufficient numbers for statistical stability and an expectation of reasonably accurate diagnosis. The purpose of age-localisation was to avoid artefacts arising from varying age distributions in different populations. For some diseases such as the congenital diseases, a different age group was more appropriate and for some a range of age groups, or

an accumulation of rates over several age groups, was used—for example, in ischaemic heart disease, cerebrovascular disease, hypertension, epilepsy, multiple sclerosis, cancer of breast, leukaemia, and cirrhosis.

Analysis proceeded as follows. First a simple correlation matrix was assembled comparing each food with each disease. For reasons given later this was repeated with Japan excluded. Correlation matrices were then produced comparing each food with every other food and each disease with every other disease. Finally, the food/disease correlations were repeated after a series of step-wise standardisation procedures.

Results

FIRST CORRELATIONS

A first matrix of 58 foods (or groupings) × 70 disease mortality sets produced 4060 correlation coefficients of which 715 (17·6%) differed from zero by ±1·96 standard errors. Repeated calculation with Japan excluded produced 610 coefficients outside this range. The main concordant results are given in Table 1 including all associations with ± r >0·78 on both occasions and >0·80 on at least one. It will be noticed that all the correlations

Table 1   *Main food disease correlations, with and without Japan*

Food × disease	Japan included	Japan excluded
Cereals × Respiratory tuberculosis	+0·80	+0·78
Sugar × Haemopoietic cancer	+0·84	+0·83
Sugar × Breast cancer	+0·84	+0·82
Meat × Colonic cancer	+0·90	+0·89
Meat × Breast cancer	+0·84	+0·78
Eggs × Colonic cancer	+0·82	+0·83
Milk × Skin cancer	+0·84	+0·82
Butter × Arthritis	+0·88	+0·88
Wine × Laryngeal cancer	+0·84	+0·84
Animal sources × Haemopoietic cancer	+0·84	+0·84
Animal sources × Breast cancer	+0·83	+0·79
Animal protein × Haemopoietic cancer	+0·86	+0·85
Animal protein × Breast cancer	+0·82	+0·78
All fats × Breast cancer	+0·91	+0·88

selected in this way were positive, and that cancer was by far the commonest disease group. In these and most subsequent calculations ± 1·96 standard errors corresponds with r = ± 0·46. Japanese mortalities are noted for high values for cancer of the stomach and cerebrovascular disease and low values for cancer of the breast and cardiovascular disease. The Japanese diet is notable for a high content of rice, of salt fish (cured or raw), and for a low level of fat intake. In fact, the strong correlation between fats and cancer of the breast was not affected by removing Japan from the analysis, and the only associations to be affected critically by

this exclusion were those of stomach cancer and cerebrovascular disease with salt fish. Nevertheless, because of the ectopic position of Japanese mortalities and diet, all subsequent analyses reported in this paper exclude Japan.

FOOD PATTERNS AND DISEASE PATTERNS

The striking feature of the full results was the systematic nature of the pattern of correlations. For many foods the same pattern of positive and negative correlations tended to recur over a range of diseases, while for other foods the sign was reversed. For example, total meat showed positive associations with multiple sclerosis, ischaemic heart disease, suicide, cancer of the large intestine, cancer of the breast, and leukaemia and so did most of the individual meats. Total cereals showed negative associations with each of these as did most of its component grains. Foods which displayed negative associations with this group of diseases tended to show positive associations with epilepsy, peptic ulcer, cirrhosis of the liver, chronic nephritis, and cancers of the larynx, mouth, oesophagus, and stomach—and vice versa. Refined sugar, milk, eggs, and butter resembled meat in their patterns of associations, while pulses, vegetables, nuts, and fish resembled the cereals. The overall pattern suggested an interaction between complex but systematic dietary and disease spectra, patterns of life-style, and geographical latitude, The question therefore arose whether within this pattern, more specific associations could be detected. A first approach was as follows.

The foods most strongly associated with (a) ischaemic heart disease, cancers of the large intestine and breast, and multiple sclerosis, or with (b) respiratory tuberculosis, peptic ulcer, cirrhosis, nephritis and carcinoma of larynx, mouth, oesophagus, and stomach were readily identified. A third class of foods showed irregular associations. An arbitrary list of 20 foodstuffs covering all three classes was then selected. These were: wheat, rice, coarse grain, pulses, vegetables, non-citrus fruits, wine, beer, sugar, potatoes, beef, pig meat, mutton, eggs, milk, cheese, fish, lard, butter, and margarine.

A matrix of correlation coefficients was next assembled, and for each food the sum of the *absolute* coefficients against all the other foods was assembled. Large sums were taken to indicate that a food would be near to one or other end of a spectrum, and small sums to indicate that it would be near the middle. Sugar and rice had the highest sums, and were negatively correlated with each other, and they were therefore used to mark the opposite ends of a scale. The other foods were then

arranged in order, to construct the list given in Table 2. The coefficients in this table indicate high levels of association between adjacent foods, and dissociations between foods situated some distance from each other. In an inexact sense this ranking evidently represents a gradient of diets passing from the Atlantic/North-Sea/Baltic coastline to the Mediterranean, from high latitudes to low, and from affluence to less than affluence. The last point is confirmed in Table 2 in the form of correlations with total calorie intakes. Those foods near the foot of the table are mainly 'primary' in the sense of being cultivated vegetables, and foods from animals not consuming cultivated crops. Those in the upper part of the table are mainly 'secondary' and therefore more expensive foods in the sense of being processed vegetable products (sugar, margarine, beer) or obtained from animals consuming cultivated crops.

Table 2 *Spectrum of foodstuffs*

	Food	Adjacent-r	Opposite-r	Total calories-r
1	Sugar	+0·48	−0·73	+0·72
2	Beer	+0·90	−0·68	+0·28
3	Pig meat	+0·58	−0·54	+0·50
4	Eggs	+0·41	−0·50	+0·45
5	Butter	+0·81	−0·54	+0·50
6	Milk	+0·35	−0·40	+0·53
7	Beef	−0·19	−0·52	+0·31
8	Margarine	+0·37	−0·41	+0·16
9	Potatoes	+0·17	+0·13	+0·30
10	Lard	+0·06	+0·06	+0·21
11	Cheese	+0·11	+0·06	+0·01
12	Mutton	−0·05	+0·13	+0·24
13	Coarse grains	+0·02	−0·41	−0·10
14	Fish	+0·24	−0·52	−0·59
15	Wine	+0·48	−0·40	−0·33
16	Fruit (non-citrus)	+0·29	−0·54	−0·62
17	Wheat	+0·45	−0·50	+0·01
18	Vegetables	+0·58	−0·54	−0·33
19	Pulses	+0·61	−0·68	−0·52
20	Rice	—	−0·73	−0·70

'Adjacent-r' is the correlation between the intake of each food and the next on the list. 'Opposite-r' is the correlation between each food, and the food symmetrically opposite in rank—for example, food 4 and food 17. 'Total calories-r' is the correlation between each food intake and the total intake of calories from all sources.

The attempted arrangement of diseases into an ordered set proved a more difficult and less meaningful task. Nevertheless, disease clusters related to common dietary patterns were evident in the data, and a selection of these patterns is given in Table 3.

STANDARDISATION

A series of standardisation procedures was undertaken in which disease mortalities were expressed as excesses or deficiencies relative to 'expected' values as determined by total calories, total fat, total protein, and others, singly and in various combinations. Some associations persisted throughout all of these procedures, while others disappeared

Table 3 *Food spectrum and certain diseases*

	Food	Tuberculosis	Stomach cancer	Peptic ulcer	Cerebro-vascular disease	Ischaemic heart disease	Cancer of large intestine	Cancer of breast
1	Sugar	−0·50	−0·42	−0·42	−0·42	0·74	0·67	0·82
2	Beer	−0·14	−0·16	−0·30	−0·43	0·11	0·66	0·70
3	Pig meat	−0·20	−0·28	−0·23	−0·42	0·30	0·57	0·67
4	Eggs	−0·34	−0·51	−0·31	−0·41	0·61	0·88	0·72
5	Butter	−0·17	−0·14	−0·18	−0·05	0·57	0·27	0·41
6	Milk	−0·25	−0·16	−0·28	−0·09	0·71	0·23	0·41
7	Beef	−0·47	−0·61	−0·40	−0·41	0·71	0·75	0·66
8	Margarine	−0·48	−0·24	−0·45	−0·47	0·14	−0·09	0·30
9	Potatoes	−0·09	0·16	0·05	0·18	0·18	0·19	0·19
10	Lard	0·28	−0·08	−0·17	−0·24	−0·18	−0·93	0·22
11	Cheese	−0·15	−0·47	−0·63	−0·55	−0·35	−0·20	0·04
12	Mutton	0·12	−0·06	−0·08	0·20	−0·04	−0·05	−0·26
13	Coarse grain	0·38	0·45	0·59	0·63	−0·16	−0·40	−0·50
14	Fish	−0·19	0·24	0·13	0·22	−0·28	−0·37	−0·44
15	Wine	0·06	0·15	0·16	0·24	−0·62	0·03	−0·34
16	Fruit (non-citrus)	0·20	0·39	0·35	0·30	−0·71	−0·37	−0·55
17	Wheat	0·78	0·23	0·25	0·26	−0·54	−0·54	−0·68
18	Vegetables	0·15	0·29	0·34	0·38	−0·64	−0·16	−0·53
19	Pulses	0·61	0·21	0·22	0·32	−0·56	−0·63	−0·80
20	Rice	0·06	0·48	0·46	0·63	−0·43	−0·45	−0·68

or appeared selectively depending upon the criterion of standardisation. These investigations were pursued on an intuitive basis relating to a number of food and disease clusters. Each will be presented separately.

ALCOHOL, CIRRHOSIS, AND DISEASES OF THE UPPER DIGESTIVE TRACT

Wine intakes were recorded in 13 European countries and varied from 8 g/day in Finland to 308 g/day in Italy and 327 g/day in France. Wine
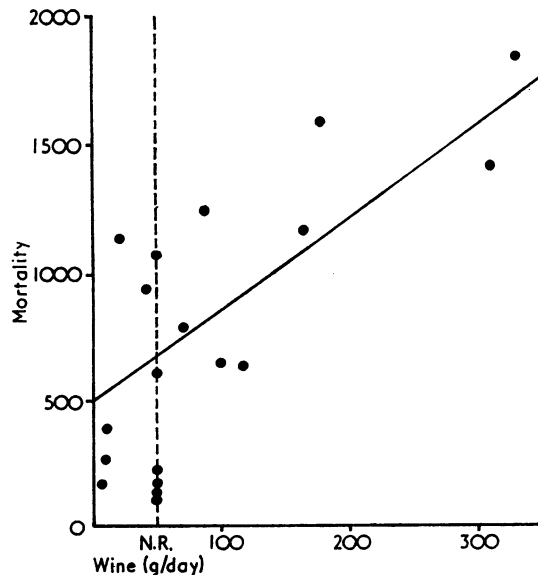
intakes were strongly associated with mortality from cirrhosis of the liver in adults, particularly in the age group 15-54 years (0·79). The association is shown in Fig. 1. Since this is a known causal effect, the strength of the statistical association provides a valuable yardstick against which other associations may be assessed when considering other causal hypotheses.

Intakes of wine also showed strong associations with mortality from cancer of the mouth (0·70) and of the larynx (0·84) (see Fig. 2) in the 55-66-year group. Cancer of the oesophagus (0·51) showed a weaker association. By contrast, cancer of the lung showed a negative correlation (−0·41), and cancer of the stomach a neutral one (0·15).

Fig. 1 *Wine intake and death from cirrhosis of liver in those aged between 15 and 64 years.*
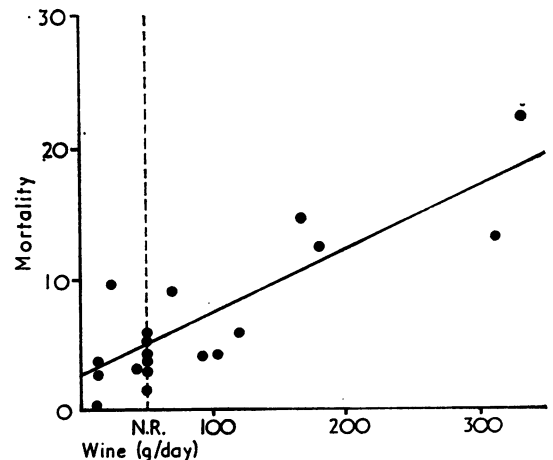

Fig. 2 *Wine intake and death from carcinoma of larynx in those aged between 55 and 64 years.*

The associations between wine and cirrhosis, cancer of the mouth and of the larynx survived a range of standardisation procedures and the association with cancer of the oesophagus was enhanced (for example, to 0·57 after standardisation for animal protein).

Intakes of beer were recorded for 10 countries and, within the limited data available, were not positively associated with cirrhosis nor with cancer of the mouth, the larynx, or oesophagus. The main unstandardised associations of beer were with multiple sclerosis (0·93), hyperplasia of the prostate (0·74), appendicitis in the age group 55-64 years (0·85), and cancer of the rectum (0·85).

GEOGRAPHICAL LATITUDE

Multiple sclerosis, suicides, arthritis, cancer of the breast, cancer of the prostate, and haemopoietic cancers were all associated with higher latitudes. Conversely, death from cirrhosis, chronic nephritis, respiratory tuberculosis, and cancer of larynx decreased at higher latitudes. The foodstuffs in the upper part of Tables 2 and 3 were eaten more in the north—particularly, sugar, potatoes, milk, margarine, and butter. Conversely, rice, vegetables, pulses, nuts, non-citrus fruits, and wine were more often eaten and drunk further south. In the present context the main interest in these latitudinal changes of mortality and diet relates to the requirement for standardisation in interpreting a wide range of positive and negative associations. Two diseases, diabetes and cancer of the skin, were however of particular interest in that their latitude and gradients were the reverse of those that might have been expected.

Diabetes, contrary to a widespread belief, did not seem to follow an affluent life-style and was more frequent in the most southern countries. It should be remembered that we are considering deaths rather than incidence, which for this disease may reflect the quality of the medical services. This is an important point which needs to be recalled in the interpretation of other north/south associations.

The other surprising result was an increase with northerliness of death from cancer of the skin, presumably both epithelioma and melanoma. Studies within individual countries have more often shown the reverse and the current material cannot be attributed to gradients similar to those suggested in explanation of the diabetes gradient. It is possible to envisage a saw-tooth north/south pattern, a southerly increase within each country running counter to an international gradient running in the other direction. For example, racial variations in skin pigmentation may more than compensate for the effects of excessive exposure to sunlight, or people accustomed to sunshine may deliberately protect themselves from it. Skin cancer may be related to episodic sunburn rather than chronic exposure. Increasing exposure to industrial carcinogens with increasing northerliness may also have contributed to a complex pattern.

VASCULAR DISEASE

Hypertensive disease, ischaemic heart disease, and cerebrovascular disease together constitute a major part of the total mortality for each country listed. Despite the common rubric of 'vascular disease' they constitute a heterogeneous group from an epidemiological point of view. Death rates from hypertensive disease and cerebrovascular disease tended to be high in the same countries and to have common foodstuff correlations, but were jointly dissociated from ischaemic heart disease mortalities and negatively correlated with its associated dietary patterns. The foodstuff associations of cerebrovascular disease and ischaemic heart disease are shown in Table 3. Ischaemic heart disease had positive associations with total animal protein (0·74), total animal fats (0·69), and total fats (0·57), while cerebrovascular disease had negative associations with these aggregates. The only notable exceptions to these inverse patterns were related to the consumption of milk, butter, cheese, and vegetable oils. Cheese and vegetable oils had negative associations with all three disease groups. Milk and butter had positive associations not only with ischaemic heart disease, but also with hypertension in the age groups 15-54, although this was not true in older age groups. These patterns withstood standardisation for total calories and for a variety of other dietary components. Because of differing theories on the aetiology of ischaemic heart disease the effects were examined of standardising for animal fats and refined sugar. In fact, the effects were barely distinguishable except that standardisation for refined sugar resulted in the deletion of a long list of previously significant associations between animal fats and a variety of diseases (not just vascular diseases), while standardisation for animal fat resulted in the deletion of a similar substantial list of unstandardised associations with refined sugar. This is scarcely surprising in view of the high positive correlation between the intakes of the two foods (0·895). The negative associations with vegetable oils and cheese persisted not only with respect to ischaemic heart disease but also for hypertension in the 15-54-year age group.

The general conclusion to be drawn is that no specific foodstuff, no foodstuff group, and none of the main calorie-providing nutrients can be

resolved on the basis of these data as a primary determinant of death rates from ischaemic heart disease, hypertension, or cerebrovascular disease.

DISEASES OF THE DIGESTIVE SYSTEM

The relationships between wine, cirrhosis, cancers of mouth, larynx, and oesophagus have been mentioned already. A more comprehensive review of digestive tract diseases shows that the pattern of food associations undergoes a systematic change as the site moves from the lips to the anus. The diseases of the upper digestive tract are associated with low calorie intakes and a high proportion of vegetable sources, and in many countries these diseases are decreasing in frequency. Those at the lower end are associated with high calorie and protein intakes and a high proportion of animal sources, and in many countries are increasing in frequency. We must recall again that in the present investigation we are concerned only with deaths, which for diseases such as appendicitis and gall bladder disease may be mainly reflections of case fatality and of the quality of available medical care, while for many of the cancers the death rates are a true reflection of incidence.

The correlation between wine and cancer of the mouth was the strongest of the associations for that disease and additional correlations with 'edible offals' (0·67), veal (0·67), and poultry (0·55) reflect the characteristic meat diets of the wine drinking countries. Wine also gave the strongest correlation with cancer of the larynx (0·84); additional associations with vegetables (0·74), veal (0·56), and nuts (0·48) can also be regarded as secondary associations. Cancer of oesophagus had weaker associations with wine intakes than did the two cancers already mentioned. Indeed, its most striking association was with intakes of horse meat (0·60). This estimate was based upon 13 countries for which statistics were quoted, and can be attributed almost entirely to the high rates and the high intakes of two of them—namely Belgium and France. Almost certainly the interpretation is indirect.

Cancer of stomach (and other gastrointestinal cancers below the stomach) showed no associations with intake of wine. Most of the food associations for stomach cancer can probably be interpreted in general rather than in specific terms as indicators of low levels of affluence. For example, its associations with the various aggregate sources of calories were all negative and the only positive associations were rather weak ones with wheat (0·52), rice (0·49), total cereals (0·48), and nuts (0·46). The association with salt/cured/dried fish (−0·27)

was negative and not significant, when Japan was excluded. It would be difficult from these data to support the suggestion made by Correa *et al.* (1975) that abrasive foods might pre-dispose.

Cancer of the large intestine excluding the rectum presented a contrasting picture. It had strong positive associations with several calorie sources, particularly animal protein (0·74), fat (0·72), and total calories (0·72), and a negative association with vegetable protein (−0·71). It is a disease of affluent northerly flesh-eating nations and the pattern of positive and negative associations is compatible with the observation that intakes of vegetable proteins were negatively correlated with intakes of animal proteins (−0·50) and with total calories (−0·78). Nevertheless, a number of individual foodstuffs and foodstuff groups yielded even stronger correlations than those of the aggregate sources including total meat (0·89), eggs (0·83), and beef (0·75). Beer (0·66), sugar (0·67), and pig meat (0·57) also supplied positive associations. The correlation with total fat is shown in Fig. 3.
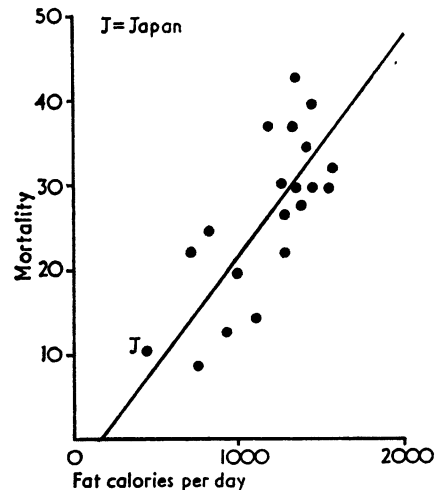

Fig. 3 *Fat intake and death from carcinoma of large intestine in those aged between 55 and 64 years.*

Cancer of the rectum behaved much like cancer of the large intestine except that most of the correlations were somewhat weaker. The strongest associations with aggregate calorie sources included total fat (0·62) and animal fat (0·60) and those with specific foodstuffs included a number containing fats such as pig meat (0·72), total meat (0·60), and eggs (0·52). Sugar (0·51) and beer (0·85) were also associated; the strength of the last association was particularly notable although based on only 10 countries.

Death from peptic ulcer followed the pattern for carcinoma of the stomach and associations with aggregate calorie sources were similar for the two diseases. Among individual foodstuffs there were only two significant positive associations with peptic ulcer—namely, total coarse grains (0·59) (including oats, barley, and maize) and nuts (0·50). The most reasonable interpretations are in general rather than in specific terms.

Deaths from gall bladder disease revealed only two significant associations—with beer (0·63) and rye (0·60). Appendicitis deaths were significantly related to the same two foods and to no others, the association with beer (0·85) being especially striking in this case. Despite the fact that gall bladder disease and appendicitis have both been linked with cancer of the large intestine and rectum as diseases determined by 'western' diets, neither—at least as far as deaths are concerned—was particularly related to intake of fat or meat.

FAT ASSOCIATED DISEASES

Associations with fat intakes have already been described both for ischaemic heart disease and cancer of the large intestine. Particularly for the second the strength of the association, the specificity of foodstuffs involved, and the resistance of the correlations to deletion through standardisation, all make for the reasonable formulation of a causal hypothesis related to total fats, particularly animal fats. The question arises whether other diseases could be incorporated within a more comprehensive postulate.

Because fat is a concentrated source of calories it is necessary to examine the pattern of both, to see if one set of associations represents a primary effect, and the other a secondary. In fact, total calories supplied weaker and fewer correlations than fats in general, or animal fats in particular. Total calories gave significant positive correlations for multiple sclerosis (0·61) and cancer of the breast (0·65). Fat intakes showed striking correlations with cancer of the breast (0·88), cancer of the haemopoietic system (0·75), multiple sclerosis (0·75), and cancer of the large intestine (0·72). There were less striking associations with cancer of the rectum (0·62), ischaemic heart disease (0·57), cancer of the prostate (0·56), and cancer of the lung (0·56). The strength of the correlation with cancer of the lung provides a useful yardstick; we know the main cause of this disease and, since it is non-dietary, we can attribute other correlations of this strength to secondary associations.

Examination was also made of division of the diet between animal and vegetable sources rather than according to nutrient content. Several diseases

followed a pattern of positive associations with animal products and negative associations with vegetable products and this must reflect the life-styles with which these diseases are associated. However, for four diseases in particular—namely, multiple sclerosis, cancer of the large intestine, cancer of the rectum, and cancer of the breast—the positive correlations with total fats were even stronger than the correlations with total animal products. The findings of chief interest are the extraordinary levels of statistical dependence upon fat of cancer of the breast (see Fig. 4), multiple sclerosis, and cancer of the large intestine (Fig. 3).
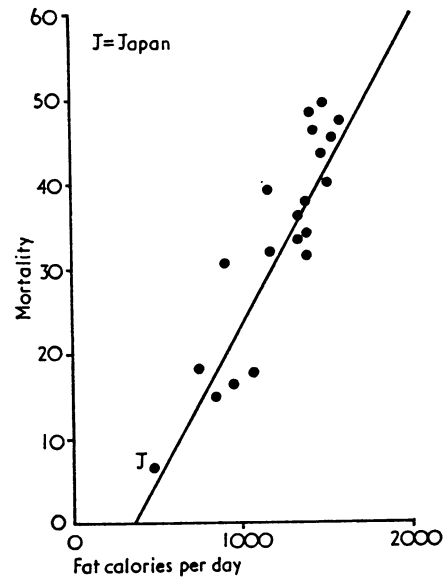

Fig. 4 *Fat intake and death from carcinoma of breast in those aged between 55 and 64 years.*

OTHER ASSOCIATIONS OF INTEREST

A number of additional associations, not already listed, raised some points of interest.

1. Nephritis and nephrosis were associated with pulses (0·78), and total cereals (0·79) including bread grain (0·69) (wheat and rye) and coarse grain (0·70) (oats, barley, and maize), and rice (0·78). In the first three respects these results resembled those for respiratory tuberculosis (0·83, 0·78, 0·61) although its associations with coarse grain (0·38) and rice (0·04) were not significant. A common interpretation relates both diseases to the diets associated with poverty, although associations between high extraction cereals

and stones in the renal tract suggest a hypothesis that may be worth exploring further.

2. Death from arthritis and spondylitis showed striking correlations with intakes of oats (0·80), sugar (0·70), milk (0·81), butter (0·88), but not with meat, eggs, or cheese. Oats and rye are atypical of cereal foods in reflecting northern patterns of living and the explanation of this association probably lies here. The extreme positions of milk and butter within this pattern, however, raises a question of interest with the possibility of a specific protein allergy.

3. Epilepsy was another disease reflecting low levels of affluence. Coarse grain (0·57), rice (0·70), and nuts (0·55) had strong positive associations, while sugar (−0·56) and eggs (−0·46) had negative ones. There were no associations, sufficiently striking to suggest specific dietary causes.

Discussion

The complexity of the associations between food-stuffs and diseases has been demonstrated. Varying reporting practices, varied precision, the different age distributions of different populations, and the concatenation of both incidence and fatality rate within mortality statistics, all give rise to difficulties. A first conclusion must be that international corre-lation analyses which limit themselves to a single disease or disease group and a small range of arbitrarily selected foods can contribute little to the formulation and testing of causal hypotheses. Some authors such as Shennan and Bishop (1974) and Armstrong and Doll (1975) have recognised these difficulties but many others have adopted too direct and naive an approach.

In these circumstances the utility of statistical procedures is difficult to assess. Significance tests provide an inefficient and a largely inappropriate screen and it may be better to use as yardsticks, estimates of correlation obtained for diseases whose aetiologies are well established, including both those where the causes are dietary and those where they are not. The correlations of cirrhosis of the liver (with alcohol), of lung cancer, road accidents, and suicide (with fat) of respiratory tuberculosis (with cereals) and of death from drowning (with fish) suggest that correlation coefficients with values less than about 0·55 are capable of being explained indirectly (if not simply as sampling variation), while those greater than about 0·70 require closer attention. The utility of standardisation procedures is also open to question;

in a complex situation the assignation of a 'primary' association to one variable and the relegation of another to 'secondary' status may depend upon quite small differences, and the outcomes may vary for different data sets. For example, Keys (1971, 1973) carried out analyses which he interpreted as showing that the relationship between athero-sclerosis and intake of sugar was secondary to an association with intake of fats. This is the reverse of Yudkin's assertion (Yudkin, 1957) and differs again from the conclusion reached in the current report—namely, that the data do not permit their differentiation. Burbank (1972) conducted a so-phisticated geographical time analysis of cancer mort-alities in the United States of America in which he identified disease 'clusters' within the total picture and attempted to relate them to identifiable and hypothetical carcinogens. This too is a more formal approach than was attempted or thought desirable in relation to the data presented here. The view taken is that sophisticated analytical methods applied to crude data should be regarded as legitimate means of display, awaiting subsequent informal interpretation, but that they are not sup-portable as formal decision processes. There is no escaping the necessity to continue to seek sets of data of different kinds, from different sources, and at different times, if we wish to discover which of the associations is sufficiently consistent to warrant attempted causal explanations.

The necessarily tentative nature of any causal hypothesis should now be evident but within these restrictions the following associations deserve special notice.

1. The strength of the correlation of wine with cirrhosis was exceeded by the strength of its association with cancer of the larynx and closely approached by that for cancer of the mouth; there was a substantial, although weaker, association with cancer of the oesophagus. The topographical continuity of these three cancer sites lends support to the hypothesis that wine drinking, or the consumption of alcoholic beverages whose amounts are highly correlated with those of wine, may contribute causally to these diseases. This is not the first time that such a suggestion has been made and, particularly for the mouth and larynx several case/control studies have demonstrated associations at the level of the individual. Kissin *et al.* (1973) demonstrated an excess of head and neck cancer in alcoholics compared with controls, while Wynder and Bross (1957), Wynder *et al.* (1957), and Vincent and Marchetta (1963) demonstrated excessive

alcohol intakes in patients with cancer of the buccal cavity, compared with controls. The last of these investigations demonstrated similar effects in relation to cancer of the pharynx and larynx, confirming a finding of Wynder *et al.* (1956) in patients with cancer of larynx. Keller and Terris (1965) conducted similar investigations in a determined effort to separate the effects of tobacco and alcohol usage, which are themselves correlated; they confirmed the conclusion of Wynder *et al.* that the effects are separate and additive. The study of Shennan and Bishop (1974) did not present or analyse information related to alcoholic drinks. Although Vogler *et al.* (1962) failed to find an association between this group of cancers and alcohol intake in the southern United States of America, the evidence of these various studies, combined with that of the current investigation can leave little doubt that wine or—as Wynder suggests—spirits, are potent causes of buccal pharyngeal and laryngeal cancers, although not exclusive causes. Clemmeson (1965) supplied an extensive review of the literature up to 1965 and Breslow and Enstrom (1974) have carried out more recent and extensive studies. These recent studies, limited to the United States of America, appeared to show that the effects of alcohol and tobacco were separable, and that both were important. This work also demonstrated the associations between intakes of beer and cancers of the rectum and colon.

2. Intakes of total fat were sufficiently highly and specifically correlated with three diseases to permit consideration of causal hypotheses. They were multiple sclerosis, cancer of the large intestine, and cancer of the breast. Cancer of the rectum had a similar but weaker pattern of association to that exhibited by cancer of the large intestine. The extent and specificity of these cancer associations was found also by Shennan and Bishop (1974) in a study of 32 countries of which 12 were also included in the current investigation. Drasar and Irving (1973) noted the high correlation between the occurrence of cancer of the colon and cancer of the breast and the high correlation between each and the intake of fat. Associations between fat intake and multiple sclerosis have been noted before; one recent study (Agranoff and Goldberg, 1974), suggested a specific association with milk

intake; this was not confirmed as an association of special note in the current investigation.

The epidemiology of multiple sclerosis drawn upon a world scale is so complex that it would be foolish to interpret the fat association in the limited context provided by these data. The disease is known to demonstrate latitudinal variation in prevalence in the same direction as variation in the fat content of diet, within this range of material; the explanation might indeed be found eventually in a component of the diet but it could also be explained in other terms and the association with diet then seen to be secondary. However, within the range of information assembled for this study it must be remarked that the correlation with fat intake is a strong one.

Cancer of the large intestine and rectum provide more direct scope for causal dietary interpretations. High fat diets probably contain relatively little fibre and to this extent the association is in keeping with current popular hypotheses (for example, Burkitt, 1971). Nevertheless, although data on fibre content were not available, there were no significant correlations between large bowel cancer and consumption of vegetables or the total grams weight of the diet. The same findings and the same comment have been made by other workers, for example, by Haenszel *et al.* (1973) in their study of large bowel cancer in immigrant Hawaiian Japanese. Several reviews of this problem (*Lancet*, 1974; *British Medical Journal*, 1975) have pointed to the weak data on which the fibre hypothesis is based and the apparently stronger case for incriminating fat among the aggregate nutrients, or of beef and sugar among the specific foodstuffs. Although the relationship between fat and breast cancer has received less attention than that between fat and colon cancer, the dependence of breast cancer mortality upon fat in the data presented here is remarkable and a causal association between this disease and this nutrient, or some component or close correlate of the nutrient, must seriously be considered.

The fat associations of breast cancer provide an indirect comment upon those with colon cancer. Within the context of a unitary causal hypothesis the findings suggest that the effect must be more than a simple mechanical process within the intestine. Unless the relationships of the two diseases represent completely different mechanisms, there is a suggestion of absorbable cancer-producing materials capable of acting at a distance. It might then be difficult, for example, to discard as indirect the fat associations of cancer of the haemopoietic system.

Reprints from: Professor E. G. Knox, Health Services Research Centre, Department of Social Medicine, Medical School, Edgbaston, Birmingham B15 2TJ.

References

Agranoff, B. W., and Goldberg, D. (1974). Diet and the geographical distribution of multiple sclerosis. *Lancet,* **2,** 1061-1066.

Armstrong, B., and Doll, R. (1975). Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *International Journal of Cancer,* **15,** 615-631.

Breslow, N. E., and Enstrom, J. E. (1974). Geographical correlations between cancer mortality rates and alcohol-tobacco consumption in the US. *Journal of the National Cancer Institute,* **53,** 631-639.

British Medical Journal (1975). Editorial: Faecal fibre fortunes. *British Medical Journal,* **2,** 580-581.

Burbank, F. (1972). A sequential space-time analysis of cancer mortality in the United States: etiologic implications. *American Journal of Epidemiology,* **95,** 393-417.

Burkitt, D. P. (1971). Some neglected leads to cancer causation. *Journal of the National Cancer Institute,* **47,** 913-919.

Clemmeson, J. (1965). Statistical studies in the aetiology of malignant neoplasms. I. Review and results. *Acta pathologica et microbiologica Scandinavica,* Supplement 174.

Correa, P., Haenszel, W., Cuello, C., Tannenbaum, S., and Archer, M. (1975). A model for gastric cancer epidemiology. *Lancet,* **2,** 58-59.

Drasar, B. S., and Irving, D. (1973). Environmental factors and cancer of the colon and breast. *British Journal of Cancer,* **27,** 167-172.

Haenszel, W., Berg, J. W., Segi, M., Kurihara, M., and Locke, F. B. (1973). Large bowel cancer in Hawaiian Japanese. *Journal of the National Cancer Institute,* **51,** 1765-1779.

Keller, A. Z., and Terris, M. (1965). The association of alcohol and tobacco with cancer of the mouth and pharynx. *American Journal of Public Health and the Nation's Health,* **55,** 1578-1585.

Keys, A. (1971). Sucrose in the diet and coronary heart disease. *Atherosclerosis,* **14,** 193-202.

Keys, A. (1973). Letter: Sucrose in the diet and coronary heart disease. *Atherosclerosis,* **18,** 352.

Kissin, B., Kaley, M. M., Su, W. H., and Lerner, R. (1973). Head and neck cancer in alcoholics. *Journal of the American Medical Association,* **224,** 1174-1175.

Lancet (1974). Editorial: Beware the ox. *Lancet,* **1,** 791-792.

Organisation for European Co-operation and Development (1970). *Food Consumption Statistics 1960-1968.* OECD: Paris.

Shennan, D. H., and Bishop, O. S. (1974). Diet and mortality from malignant disease in 32 countries. *West Indian Medical Journal,* **23,** 44-53.

Vincent, R. G., and Marchetta, F. (1963). The relationship of the use of tobacco and alcohol to cancer of the oral cavity, pharynx, and larynx. *American Journal of Surgery,* **106,** 501-505.

Vogler, W. R., Lloyd, J. W., and Milmore, B. K. (1962). A retrospective study of etiological factors in cancer of the mouth, pharynx, and larynx. *Cancer,* **15,** 246-258.

World Health Organisation (1967). *International Classification of Diseases,* 8th revision. WHO: Geneva.

World Health Organisation (1973). Vital statistics and causes of death. In *World Health Annual Statistics for 1970.* WHO: Geneva.

Wynder, E. L., and Bross, I. J. (1957). Aetiological factors in mouth cancer. *British Medical Journal,* **1,** 1137-1143.

Wynder, E. L., Bross, I. J., and Day, E. (1956). A study of environmental factors in cancer of the larynx. *Cancer,* **9,** 86-110.

Wynder, E. L., Bross, I. J., and Feldman, R. M. (1957). A study of the etiological factors in cancer of the mouth. *Cancer,* **10,** 1300-1323.

Yudkin, J. (1957). Diet and coronary thrombosis: hypothesis and fact. *Lancet,* **2,** 155-162.