SCIENTIFIC ARTICLES

Fat in the Diet and Mortality from Heart Disease

A Methodologic Note

J. YERUSHALMY, PH.D., BERKELEY, CALIFORNIA, AND HERMAN E. HILLEBOE, M.D., ALBANY, NEW YORK

(From the Division of Biostatistics, School of Public Health, University of California, Berkeley, California, and the New York State Department of Health, Albany, New York

THE growing importance of heart diseases as a cause of disability and death has stimulated widespread investigation into possible causative factors, especially those of environmental origin. Within the last few years certain components of the diet, especially fat, have attracted increasing notice as suspected causal agents. The evidence which has been presented for the existence of a relationship between diet and heart disease, however, is, for the most part, derived from indirect methods of study. In these indirect methods the primary unit of observation is the *group;* in the direct method the primary unit of observation is the *individual.* It is a common practice to use indirect methods to construct indices of the dietary habits of national, racial, social, or other special groups, and to relate these indices to the frequencies of various forms of heart disease in these groups. The data in many cases are based on vital statistics and on statistics concerning the food supplies available for consumption in different countries.

It is well known that the indirect method merely suggests that there is an *association* between the characteristics studied and mortality rates and, further, that no matter how plausible such an association may appear, it is not in itself proof of a cause-effect relationship. But quotation and repetition of the suggestive association soon creates the impression that the relationship is truly valid, and ultimately it acquires status as a supporting link in a chain of presumed proof.

There is no question that indirect methods of study are valuable when properly employed, especially as indicating suitable paths for further investigation. Thus, morbidity and mortality from certain respiratory diseases have been compared in urban and rural areas for clues as to the role of air pollution; the trend of lung cancer mortality has been investigated in relation to cigaret consumption, and death rates from specific diseases have been examined in census-tracts which are distinguished by measurable environmental characteristics. Indeed, so complex are current health problems that associations between various characteristics of the environment and morbidity and mortality should be more intensively studied; the indirect method is often the only practicable procedure.

Unfortunately, however, if used superficially and not properly augmented, the indirect method has many weaknesses. The most serious is the fact that the apparent association often proves to be the result of nonpertinent extraneous factors. Therefore, it is always necessary to probe further, to go beyond the simple, apparent association and to investigate related variables. This does not mean that the worker must undertake exhaustive studies to rule out all environmental factors other than the suspected one. In many cases systematic testing of a few related environmental factors

Presented at a meeting of the American Epidemiological Society, Albany, New York, April 5, 1957.

against one or two additional disease entities will disclose whether the presumed association is valid.

In this presentation an attempt is made, first, to determine the sort of supplementary investigation which can make the indirect method a more dependable tool for studying the association between disease entities and environmental factors and, second, to determine the validity of the presumed association between national death rates from heart disease and national dietary-fat levels.

An actual example illustrates how the impression that there is a strong association between mortality from heart disease and proportion of fat available in the diet in different countries assumed the stature of a proved fact.* In 1953 Keys[1] published a chart (Fig. 1) which shows the relationship in six different countries between the national death rate for men aged forty-five to forty-nine and fifty-five to fifty-nine from "arteriosclerotic and degenerative heart disease" and the proportion of fat-calories available in the respective national diets. Taken by itself, the figure is impressive; there is regular progression and concurrent variability of the two factors. The author cautiously observed that "it must be concluded that dietary fat somehow is associated with cardiac disease mortality at least in middle age."

In 1954 Leitner[2] reproduced this figure in *The Lancet*, but his interpretation of it gave the association greater stature and definitiveness: "There appears to be a strong, if not convincing correlation between the amount of fat in the diet and the death rate from degenerative heart disease."

In 1955 in another publication Keys[3] referred to the same evidence from his 1953 figure, but this time he interpreted the data in much stronger terms: "The analysis of international vital statistics shows a striking feature when the national food consumption statistics are studied in parallel. Then it appears that for men aged forty to sixty or seventy, that is, at the ages when the fatal results of atherosclerosis are most prominent, there is a remarkable relationship between the death rate from degenerative heart disease and the proportion of fat calories in the national diet. A regular progression exists from Japan through Italy, Sweden, England and Wales, Canada and Australia to the United States. No other variable in the mode of life besides the fat calories in the diet is known which shows anything like such a consistent relationship to the mortality rate from coronary or degenerative heart disease."

Clearly it has become important to determine whether these statements reflect the known or ascertainable facts. But to do so it is essential to consider questions of methodology, namely, the elementary operations and procedures by which it is possible to determine whether the association is valid.

The term "valid" association, as used in this discussion, does not necessarily denote a cause-effect relationship; rather it signifies that an association is "specific." That is, validity

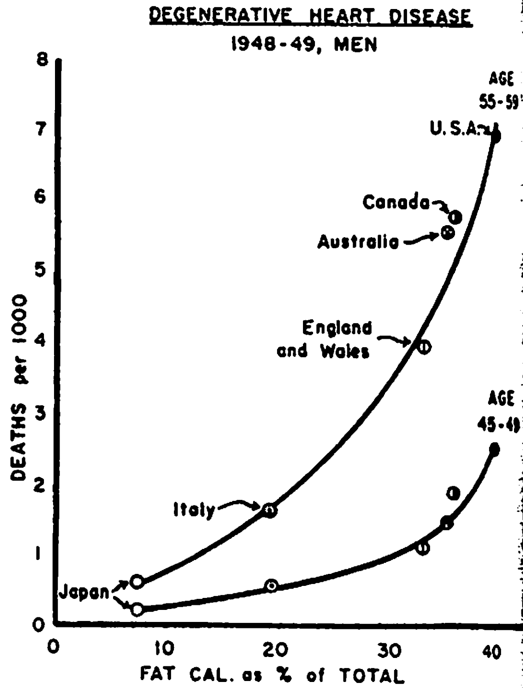


FIG. 1. Mortality from degenerative heart disease (Categories 93 and 94 in the Revision of 1938, categories 420 and 422 in the Revision of 1948, International List). National vital statistics from official sources. Fat calories as percentage of total calories calculated from national food balance data for 1948 supplied by the Nutrition Division, Food and Agriculture Organization of the United Nations. (Illustration and legend taken from Keys,[1] courtesy of *Journal of Mount Sinai Hospital*)

* The authors became interested in this question during a meeting of a WHO study group on atherosclerosis and ischemic heart disease in Geneva, Switzerland, in November, 1955. At that meeting statements were made about the association between heart disease mortality and fat in the diet. As it later developed, these were based on a few selected countries and were of questionable validity. On their return to the United States the authors reviewed the available data carefully, and the results indicated that the subject required further study.

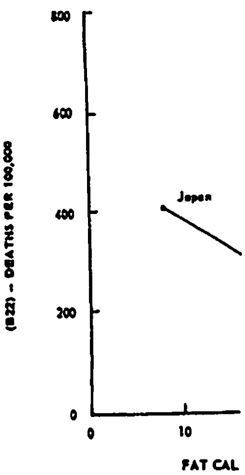


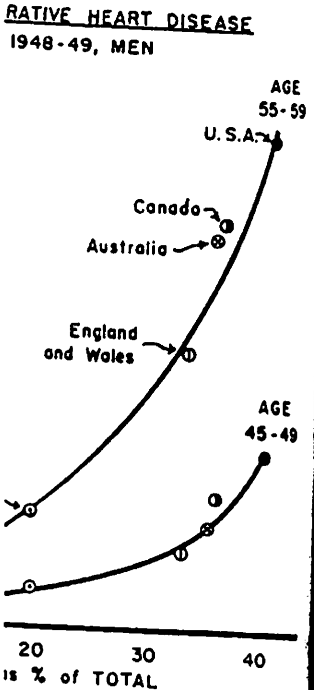
FIG. 2. Mortality fr
the central nervous syst
per cent of total calories i
years from six countries
pose. Calculated from r
F.A.O. (see text for defin

implies that the asso
ables is in fact betw
gated and does not r
with a broader group,
of the variables forms
tigator must search
whether there is a s
related factor is subs
suspicion or when oth
stituted for the disea

This paper descril
dures by which valid
procedures involve
the data from which
structed and two tes
present case this in
association between
components *other tha*
between dietary fat a
diseases of the heart.

Preliminary Eval

Before undertaki
evaluating the "spe
sociation, it is ofte
methods by which t
selected, the definit
data, and the amc
pecially in "group
the dietary fat-hear
preliminary steps ar



ɔm degenerative heart disease
n the Revision of 1938, cate-
the Revision of 1948, Inter-
ı vital statistics from official
ı percentage of total calories
. food balance data for 1949
ɔn Division, Food and Agri-
the United Nations. (Illus-
ɜn from Keys,[1] courtesy of
Hospital)

is a remarkable relation-
h rate from degenerative
proportion of fat calories
A regular progression
through Italy, Sweden,
Canada and Australia to
o other variable in the
e fat calories in the diet
anything like such a con-
the mortality rate from
heart disease."
important to determine
ts reflect the known or
t to do so it is essential
f methodology, namely,
ons and procedures by
determine whether the

ciation, as used in this
essarily denote a cause-
er it signifies that an
." That is, validity

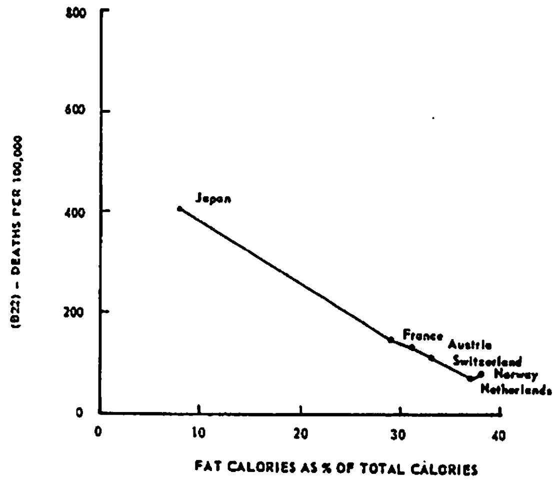


Fig. 2. Mortality from vascular lesions affecting the central nervous system (B-22) and fat calories as per cent of total calories in males fifty-five to fifty-nine years from six countries selected for this specific purpose. Calculated from national food balance data by F.A.O. (see text for definition).

implies that the association between two variables is in fact between the variables investigated and does not merely reflect relationships with a broader group, of which one or the other of the variables forms a part. Hence, the investigator must search for ways to determine whether there is a similar association when a related factor is substituted for the one under suspicion or when other disease entities are substituted for the disease being studied.

This paper describes the elementary procedures by which validity may be tested. These procedures involve preliminary evaluation of the data from which the indices have been constructed and two tests for "specificity." In the present case this involves examining: (1) the association between heart disease and dietary components *other than fat* and (2) the association between dietary fat and causes of death *other than diseases of the heart.*

Preliminary Evaluation of Data

Before undertaking the essential task of evaluating the "specificity" of a seeming association, it is often necessary to examine the methods by which the primary data have been selected, the definitions and limitations of the data, and the amount of detail required, especially in "grouping" causes of death. In the dietary fat-heart disease problem all of the preliminary steps are important.

Selection of Data.—In studies of association, as in any other study, the method of selecting the data largely determines whether the results can safely be generalized. For example, the association shown in Fig. 1 and the conclusion drawn from it by its author convey the impression that it reflects a general phenomenon. Hence, it is of first importance to know on what basis the six countries were selected, in order to determine whether the findings for them can be generalized to other countries.*

Since no information is given by Keys on how or why the six countries were selected for Fig. 1, it is necessary to investigate the association between dietary fat and heart disease mortality in *all* countries for which information is available. This is shown in Fig. 3 for males aged fifty-five to fifty-nine years in 22 countries.† It is immediately obvious that the inclusion of all the countries greatly reduces the apparent association.

For example, it may be seen from Fig. 3 that in the narrow band between 30 per cent and 40 per cent dietary fat, there appears the entire gamut of heart disease mortality, ranging from less than 300 per 100,000 for Austria, West Germany, Sweden, Norway, Denmark, and the Netherlands, to 600 or more for Australia, Canada, and Finland, and as much as 739 for the United States. Indeed, within this higher dietary fat band seven of the countries have 400 or fewer deaths from heart disease per 100,000, while six have rates of more than 400. Thus, we see that in countries with approximately the same proportion of the diet available as fat, the

* As an illustration of the importance of this question of selection, it may be of interest to show how by selection of countries it is possible to "demonstrate" an apparent inverse relationship between per cent of calories from fat and mortality from vascular lesions affecting the central nervous system (B-22 in the Abbreviated International List). These data are presented for six countries in Fig. 2. If Fig. 1 is taken as supporting evidence for a direct association between dietary fat and degenerative heart disease mortality, on the same basis one may take Fig. 2 as indicating an antagonism between fat in the diet and death rates from cerebral hemorrhage. Of course it is obvious that no such conclusion should be inferred. Parenthetically it is interesting to note that Japan, which has the lowest death rate from cardiovascular disease, has by far the highest death rate from cerebrovascular disease.

† Age group fifty-five to fifty-nine is used in this and subsequent figures and tables because the association shown in Fig. 1 is more striking for this age group. However, almost identical findings can be shown for other age groups or for age-adjusted rates. The detailed data for such an analysis are given in Tables I and II of Appendix A. Data are presented for all countries for which the dietary source of calories was provided by the Food and Agriculture Administration of the United Nations[4] and for which deaths from disease of the heart and recent population estimates are available.[5] There were 22 countries which met these criteria.

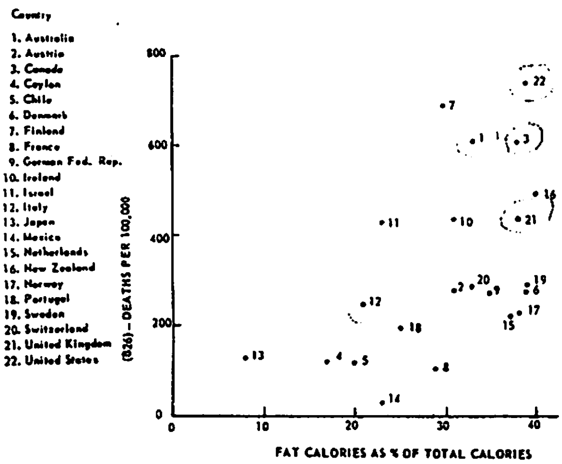


FIG. 3. Mortality from arteriosclerotic and degenerative heart disease (B-26) and fat calories as per cent of total calories in males fifty-five to fifty-nine years. Calculated from national food balance data by F.A.O. (see text for definition).

heart-disease mortality ranges from 220 to 739 per 100,000.

Hence, the selection of the original six countries, for whatever reason, greatly exaggerated the importance of the association. Nevertheless, the scatter of points in Fig. 3 suggests that there is some association in the conventional sense between the two variables; it is therefore necessary, first, to examine the basic data further and, second, to determine whether the association is "specific" for the two variables.

DEFINITIONS AND LIMITATIONS OF DATA.—In studies of association involving indices of environmental characteristics and mortality from specific diseases, it is well to keep in mind that both these data are subject to great variations among countries. For example, the index for fat as a percentage of total calories in the diet used in the original study quoted[1] is not based on a knowledge of *food consumption* in the various countries but at best represents the relative amounts *available for consumption*. These indices were constructed by the Food and Agriculture Organization of the United Nations from statistics on production, imports, exports, and on the proportion of available food used for purposes other than human nutrition. The underlying data are stated by F.A.O. to be subject to great limitations (see Appendix B). Moreover, there are no doubt great differentials in food "scraps" in the various countries being compared. For example, it is highly probable that far more edible dietary fat is thrown into waste cans in the United States than in less fortunate countries.

The mortality data, which are derived from death certificates and population estimates, are similarly subject to limitations and great variations from country to country. Accuracy and completeness of reports are dependent on the quantity and quality of medical care, on local customs, and on the diagnostic habits of the medical profession. It is necessary, therefore, to bear in mind that association between such variables may be only apparent and may reflect the differences in definition and reporting patterns among the countries under study.

GROUPINGS OF "CAUSES OF DEATH."—In investigations involving vital statistics, especially those of different countries, a problem arises as to the best method of "grouping" the individual disease entities into categories and subcategories of causes of death. Because of variations from country to country in diagnostic acumen and facilities and in cultural patterns, deaths due to a specific disease are not always similarly designated and recorded. Often, however, they may be recorded under related titles. Consequently, grouping into broader categories will diminish the resultant variability in many instances. It is not easy to determine in each situation the grouping which provides the most sensitive measure for a specific purpose. There are, however, certain elementary considerations which may be helpful.

Often a comparison of the distribution of the most specific and exclusive group with that of the less definitive group provides a basis for judgment. This may be illustrated by attempting to determine the appropriate groupings of diseases of the heart for the purpose of studying their association with dietary fat.

In most countries deaths are classified according to the International List of Causes of Death; for international comparisons an abbreviated list is generally used. The latter divides diseases of the heart into the following fairly broad groups:

B-24 Rheumatic fever
B-25 Chronic rheumatic heart disease
B-26 Arteriosclerotic and degenerative heart disease
B-27 Other diseases of the heart
B-28 Hypertension with heart disease

Of these B-26 appears to be the group most specifically relevant to the study at hand. How-

hrown into waste cans in
in less fortunate countries.
, which are derived from
population estimates, are
mitations and great vari-
country. Accuracy and
ts are dependent on the
of medical care, on local
diagnostic habits of the
t is necessary, therefore,
association between such
apparent and may reflect
ion and reporting patterns
der study.

AUSES OF DEATH."—In
vital statistics, especially
tries, a problem arises as
'grouping" the individual
egories and subcategories
ecause of variations from
diagnostic acumen and
patterns, deaths due to a
always similarly desig-
ften, however, they may
ed titles. Consequently,
categories will diminish
in many instances. It
ne in each situation the
les the most sensitive
purpose. There are,
ary considerations which

f the distribution of the
ive group with that of
p provides a basis for
illustrated by attempt-
ppropriate groupings of
the purpose of studying
ary fat.

hs are classified accord-
ist of Causes of Death;
risons an abbreviated
e latter divides diseases
following fairly broad

ic heart disease
and degenerative heart

the heart
h heart disease

to be the group most
study at hand. How-

TABLE I.—SUBCATEGORIES* OF THE GROUP "ARTERIOSCLEROTIC AND DEGENERATIVE HEART DISEASES" (CATEGORY B-26)†

Country	Per Cent of All B-26 Deaths in the Subcategories 420	421	422	Rates per 100,000 420	421	422
Australia	88.7	2.4	8.9	516	14	52
Canada	93.5	0.7	5.8	550	4	34
England and Wales	87.1	3.3	9.6	372	14	41
Italy	48.9	7.8	43.3	120	19	106
Japan	41.1	41.4	17.5	50	51	21
United States	93.7	1.2	5.1	660	8	36

* Explanations of subcategories: 420—arteriosclerotic and coronary disease; 421—chronic endocarditis; 422—other myocardial degeneration.
† Males 55 to 59 years in six countries used by Keys.

ever, in the detailed International List B-26 is further divided into the following subgroups:

B-26:
- 420 Arteriosclerotic heart disease, including coronary disease
- 421 Chronic endocarditis not specified as rheumatic
- 422 Other myocardial degeneration

According to this list the most specific and relevant group appears to be the one composed of subgroups 420 and 422. Indeed, this is the group used by Keys for the data presented in Fig. 1.

The question is whether this group would provide the most sensitive measure for the purpose. This can be answered by comparing the distribution of the rates in the less definitive subgroup with that of the most specific and exclusive subgroup. The following general principle often may be found useful: If the distribution of the rates in the less definitive subgroup is not unusual, i.e., if the variability by country is not extreme, then little will be gained and much may be lost from the inclusion of the less definitive subgroup. However, if the distribution is unusual, so that a few countries have rates in the less definitive subgroup which are either extraordinarily high or low—and especially if these unusual rates are opposite in direction to those in the more specific and exclusive subgroup—then it is probably safer to include the nondefinitive subgroup with the more precise ones. At least it is necessary to provide a plausible explanation for the extreme rates in the few countries.

The application of this principle to the above problem will help to clarify this point. In deciding whether to include the subgroup 421, which from the standpoint of pathology may not even belong in the group of arteriosclerotic

TABLE II.—MORTALITY FROM "OTHER DISEASES OF THE HEART" (CATEGORY B-27)*

COUNTRIES RANKED ACCORDING TO RATES AMONG MALES IN THE AGE GROUP 55 TO 59 YEARS (1951–1953). CORRESPONDING RATES FOR DEATHS CLASSIFIED AS B-26 ARE GIVEN FOR COMPARISON

Rank	Country	Rates per 100,000 B-27	B-26
1	Chile	488	115
2	Mexico	241	28
3	France	180	102
4	Austria	77	279
5	Finland	68	689
6	Ireland	55	435
..	...	...	...
..	...	...	...
..	...	...	...
..	...	...	...
11	Switzerland	44	287
..	...	...	...
13	Germany	40	276
14	Italy	40	428
..	...	...	...
16	United States	35	739
..	...	...	...
..	...	...	...
..	...	...	...
20	Canada	23	609
21	Netherlands	23	220
22	United Kingdom	21	434

* Source: WHO, *Annual Epidemiological and Vital Statistics, 1951–1953.*

heart disease,* the distribution of the rate for this subgroup in the six countries should be reviewed.[6] It is seen from Table I that the distribution is not reasonably uniform in all six countries, both as to rates per 100,000 and as to percentages of all B-26 deaths. Japan reports a death rate for the subgroup 421 strikingly higher than that of any of the other countries and at the same time reports a death rate for the subgroup 420 which is strikingly lower. In the absence of a plausible explanation for the unusually high rates for chronic endocarditis in Japan alone, it is clearly appropriate to include 421 and consider the entire group of B-26.

Similarly it is necessary to proceed in the same way and to decide whether B-27 ("other diseases of the heart") should be added to B-26 to form a combined group for purposes of the studies on dietary fat. From Table II it is seen that in this case too the distribution of the death rates from the less definitive group (B-27) in the 22 countries is very unusual; a few countries stand out with rates which are far out of line with the other countries. As compared with a mean death rate for B-27 of 56.7 per 100,000 for all 22 countries and a median rate of 44, France reports a rate of 180, Mexico 241, and Chile 488.

* Incidentally, it becomes almost mandatory to include 421 with the other components of B-26 in studies of this kind because much of the published statistics on an international level are provided only by the abbreviated list.

TABLE III.—Rank Correlation Coefficients Between Various Dietary Components and Death Rates for Different Groupings of Categories* of Diseases of the Heart in 22 Countries (Males 55 to 59 Years)

Dietary Components†	Groupings of Categories of Diseases of the Heart: B-26	B-26 + B-27	B-26 + B-27 + B-28	B-25 + B-26 + B-27 + B-28
Number of calories				
Total calories	0.723	0.593	0.619	0.637
Calories from fat	0.659	0.470	0.508	0.523
Animal fat (N = 21)**	0.684	0.562	0.610	0.604
Vegetable fat (N = 21)**	−0.236	−0.282	−0.187	−0.186
Calories from protein	0.709	0.694	0.691	0.692
Animal protein	0.756	0.695	0.708	0.708
Vegetable protein	−0.430	−0.153	−0.197	−0.181
Calories from carbohydrate	0.305	0.423	0.390	0.414
Per cent of total calories				
From fat	0.587	0.390	0.426	0.436
Animal fat (N = 21)**	0.677	0.557	0.610	0.604
Vegetable fat (N = 21)**	−0.468	−0.509	−0.526	−0.531
From protein	0.172	0.465	0.421	0.411
Animal protein	0.643	0.608	0.616	0.608
Vegetable protein	−0.651	−0.483	−0.519	−0.411
From carbohydrate	−0.562	−0.386	−0.415	−0.423

Critical Values of r

N	$\alpha = 0.05$	$\alpha = 0.02$
21	±0.438	±0.521
22	±0.428	±0.508

* Categories of diseases: B-25—chronic rheumatic heart disease; B-26—arteriosclerotic and degenerative heart disease; B-27—other diseases of the heart; B-28—hypertension with heart disease.
† Calculated from national food-balance data by F.A.O. (see text for definition).
** Data not available for France.

On the other hand, while the mean death rate for B-26 for the 22 countries is 411.6 and the median rate is 278, Chile reports a rate of only 115, France 102, and Mexico the startlingly low rate of 28. Here also, unless there is a reasonable explanation for the high rates in these countries in this less definitive group of "other diseases of the heart," it may be safer to operate on the assumption that in these three countries some deaths from arteriosclerotic and degenerative heart disease are being recorded under the broad group of "other diseases of the heart." Consequently the group B-27 should be included with B-26 to form a combined group for the purposes of studying its correlation with components of the diets in the different countries.

When the groups B-25 and B-28 were analyzed in the same way, no unusual distributions appeared. Their inclusion would gain little and may reduce the sensitivity of the index; therefore, they should not be included.

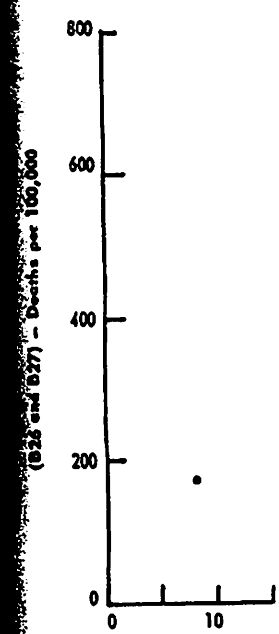
It may be concluded then that for investigating the association between dietary fat and mortality from arteriosclerotic and degenerative heart disease, the most appropriate group is one that is composed of B-26 and B-27. Accordingly this grouping is used in most of the figures and tables in this paper. In Table III however, data from all the groups B-25 to B-28 are presented in several combinations in order that the readermay judge for himself the effects of various groupings.

This latter procedure may be a useful one to adopt; if there is any doubt as to the most appropriate grouping, it is advisable to present the data in various groupings, as is done in Table III, taking care to include that grouping which lends least support to the hypothesis under consideration.

Tests for "Specificity" of an Association

The crucial tests are those for "specificity" to determine whether the suggestive association is specific for the variables themselves or merely reflects a diffuse relationship with other factors. The tests for specificity involve examining each of the variables in turn for possible associations with factors related to the other variable.

Thus, in the association considered in this paper it is necessary to test, first, whether the association between heart disease mortality and diet holds true only for fat or whether one may find a similar association between heart disease and other components of the diet, such as protein or carbohydrate. In more general terms the procedure may be stated as follows:
When a single component (fat) of an environmental characteristic (diet) is suspected of being associated with a disease entity, it is necessary to determine whether there is a similar association with other



FAT CAL
OF TO

FIG. 4. Mortality from
(B-26 and B-27) and fat calo
calories (Panel II). Calcul

components of the charact
Similarly it is necessa
associated only with art
ative heart disease or
sociated with other di
general terms the proce
If a component (fat) of
suspected of being ass
entity (a specified group
it must be seen, in ad
associations between th
characteristic and other c
It is clear that if re
characteristic are also
disease and moreover if
is associated also with
ponent is unlikely to be
with the disease under
FAT VERSUS OTHE
DIET.—In Fig. 4 deal
the heart (B-26 and B
plotted against the pe
(Panel I) and from a
available for consump
of the scatter of points

ts and Death Rates for Dif-
(Males 55 to 59 Years)

es of the Heart—
+ B-25 + B-26 + B-27 + B-28
0.637
0.523
0.604
−0.186
0.692
0.708
−0.181
0.414
0.436
0.604
−0.531
0.411
0.608
−0.411
−0.423

Critical Values of r

N	α = 0.05	α = 0.02
21	±0.438	±0.521
22	±0.428	±0.508

ınd degenerative heart disease;

may be a useful one to
· doubt as to the most
; is advisable to present
pings, as is done in Table
ıde that grouping which
the hypothesis under

y" of an

those for "specificity"
ıe suggestive association
les themselves or merely
ıship with other factors.
ıvolve examining each of
or possible associations
: other variable.
ion considered in this
test, first, whether the
t disease mortality and
at or whether one may
ı between heart disease
the diet, such as protein
ore general terms the
! as follows:
fat) of an environmental
pected of being associated
s necessary to determine
r association with other

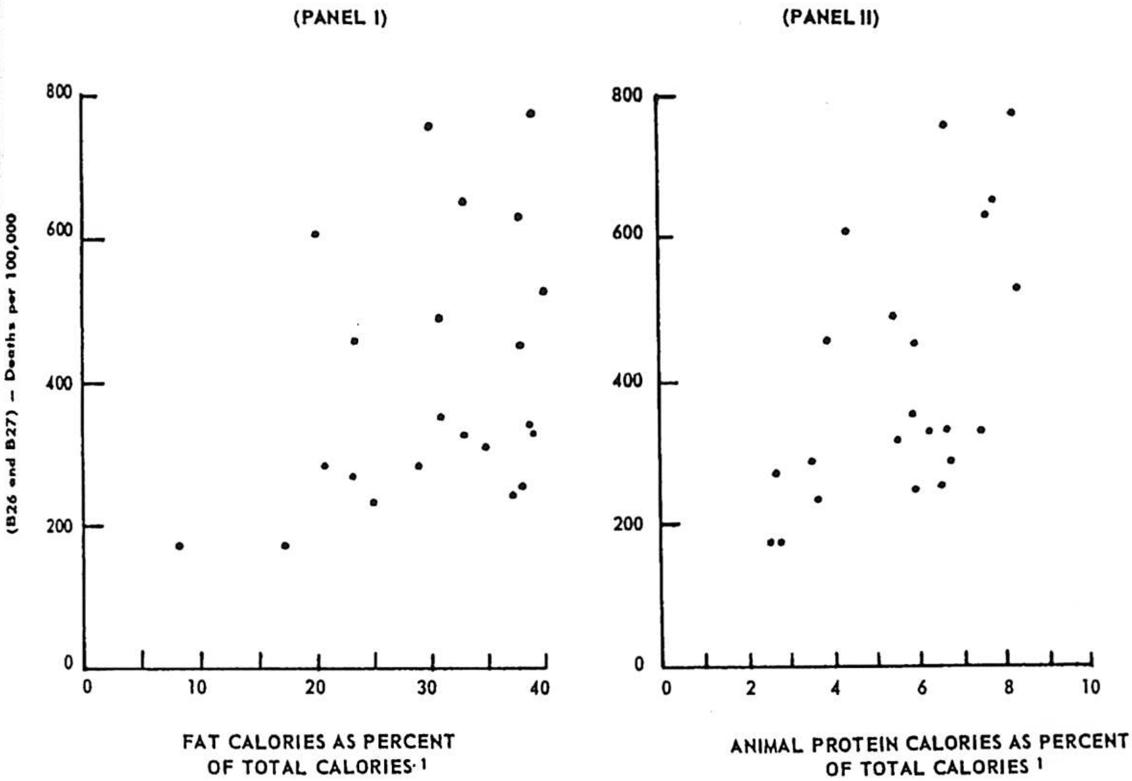


Fig. 4. Mortality from "arteriosclerotic and degenerative heart diseases" and "other diseases of the heart" (B-26 and B-27) and fat calories as per cent of total calories (Panel I) and animal protein calories as per cent of total calories (Panel II). Calculated from national food balance data by F.A.O. (see text for definition).

components of the characteristic.

Similarly it is necessary to test whether fat is associated only with arteriosclerotic and degenerative heart disease or whether fat may be associated with other disease entities. In more general terms the procedure may be stated thus: *If a component (fat) of a characteristic (diet) is suspected of being associated with one disease entity (a specified group of diseases of the heart), it must be seen, in addition, whether there are associations between the same component of the characteristic and other disease entities.*

It is clear that if related components of the characteristic are also associated with the same disease and moreover if the suspected component is associated also with other diseases, the component is unlikely to be "specifically" associated with the disease under study.

Fat Versus Other Components of the Diet.—In Fig. 4 death rates from diseases of the heart (B-26 and B-27) for 22 countries are plotted against the per cent of calories from *fat* (Panel I) and from *animal protein* (Panel II) available for consumption. Visual comparison of the scatter of points in the two panels indicates that the association between dietary protein and heart disease (Panel II) is at least as strong as that between dietary fat and heart disease (Panel I). This clearly suggests that the dietary fat-heart disease association is *not* unique or specific since the association between *fat* and heart disease mortality is not as strong as that between *animal protein* and heart disease.

The associations between heart disease mortality and various components of the diet can be evaluated by calculating the rank correlation coefficients of the variables.[7] In Table III these coefficients for various components of the diet and various groupings of diseases of the heart are presented. These data provide a measure of the strength of association between any of the variables. Analysis of these coefficients shows that whether measured in absolute number of calories or as per cent of calories, available dietary fat is not more strongly associated with heart disease mortality than are the other components of the diet. Indeed, for the combined grouping B-26 and B-27, shown above to be most appropriate for this study, the rank correlation coefficient with dietary fat as percentage of total calories is not significant at either the 2 per cent or 5 per cent level, whereas that for protein is significant at both levels. In all categories of heart disease the association is strongest for animal protein expressed in total calories.

The rank correlation coefficient for the association between per cent of calories from fat and heart disease mortality is highest (0.587) for B-26 alone, but even in this group the coefficient for per cent of animal protein and heart disease mortality is higher (0.643). In each of the heart disease groupings the strongest association is with total number of calories from animal protein, and in each case the rank correlation coefficient is statistically significant.

Nevertheless, it must be emphatically stated that the authors do not intend to suggest that the association between dietary protein and heart disease is valid. That is a complex subject which in itself is worthy of further independent investigation. The data are presented to illustrate the steps necessary to test the "specificity" of the dietary fat-heart disease association. It can be said that the data in Table III alone do *not* support the view that dietary fat is uniquely or unequivocally associated with mortality from heart disease.

On the other hand, these data do not at this point *rule out* the possibility. It may be that fat and animal protein are each related to heart disease mortality; it is now necessary to test whether either of these components is specifically associated with heart disease or whether either is associated with noncardiac diseases as well.

HEART DISEASE VERSUS OTHER CAUSES OF DEATH.—In Table IV the rank correlation coefficients for various components of the diet with death rates from noncardiac causes and with death rates from all causes are presented. Table IV shows that fat calories and animal protein calories, which were seen above to be positively associated with heart disease, are here *negatively** associated with noncardiac diseases. It would be difficult to explain these negative correlation coefficients as valid expressions of antagonistic tendencies between animal protein and fat per se and mortality from all causes other than heart disease. A more plausible explanation

TABLE IV.—RANK CORRELATION COEFFICIENTS BETWE[EN] VARIOUS DIETARY COMPONENTS AND DEATH RATES [FROM] ALL CAUSES AND FROM ALL CAUSES OTHER THAN DISEA[SES] OF THE HEART (B-25 TO B-28) AND PER CENT OF DEA[THS] ASSIGNED TO SENILITY, ILL-DEFINED, AND UNKNOWN CAU[SES] (B-45)*
(Males 55–59 Years)

Dietary Components†	All Causes	All Causes Except B-25, B-26, B-27, B-28	Deaths in Category B-4[5] As Percent [of] All Deaths
Number of calories			
Total calories	−0.189	−0.530	−0.549
Calories from fat	−0.340	−0.674	−0.433
Animal fat (N = 21)**	−0.169	−0.466	−0.592
Vegetable fat (N = 22)**	−0.339	0.296	0.461
Calories from protein	−0.067	−0.398	−0.487
Animal protein	−0.099	−0.505	−0.572
Vegetable protein	0.275	0.452	0.200
Calories from carbohydrate	0.294	0.172	−0.469
Per cent of total calories			
From fat	−0.372	−0.657	−0.375
Animal fat (N = 21)**	−0.178	−0.481	−0.592
Vegetable fat (N = 22)**	−0.277	−0.090	0.632
From protein	0.235	−0.086	−0.275
Animal protein	−0.106	−0.405	−0.479
Vegetable protein	0.187	0.521	0.442
From carbohydrate	0.396	0.671	0.353

CRITICAL VALUES OF r

N	$\alpha = 0.05$	$\alpha = 0.0[2]$
21	±0.438	±0.521
22	±0.428	±0.508

* Males 55 to 59 years.
† Calculated from national food balance data by F.A.O. (see text for definition).
** Data not available for France.

is that the dietary components which according to the rank correlation coefficients appeared to be positively related to heart disease are indices which reflect related attributes of the various countries. That is, it may be that the amount of fat and protein available for consumption is an index of a country's development, industrially, nutritionally, medically, and no doubt in other respects as well. Some of these environmental factors themselves may be related to heart disease mortality, or in fact, there may be an association between heart disease and animal protein or fat (although neither association has been shown to be valid). It may also be that countries with more abundant diet are more highly developed and diagnostic acumen is greater. Hence, it is possible that in some of the countries in which less protein and fat are available, a certain percentage of the deaths from

* If a rank correlation coefficient is positive, the association is direct; if negative, then the variables are inversely related—that is, high values for one variable tend to be associated with low values for the other.

arteriosclerotic and de
are recorded under the
This possibility is s
by the fact that Table
ciation between dietar
tality from "all causes
of the heart form su
deaths in the age gro
if there is any associat
ponents and heart dis
show up also in their
of death. That ther
again suggests that t
disease is not specific.
shows, there are appre
coefficients between
death rates from B-45
unknown causes"). T
considered a rough ind
of death certification
The negative associat
is further evidence of t
presumed association.

In summary, the e
for which data are a
association between th
available for consump
and mortality from
erative heart disease i
is specific neither for
disease mortality. C
ciation cannot serve
hypothesis which im
factor in arterioscler
disease.

Comment

The problem of id
chronic diseases of u
ingly complex and ca
or decisive experime
of evidence must b
analyzed, and tested
ships before an etiolo
The "indirect method
ble tool in this field
erly and in conjuncti

The basic material
comprises statistics
or morbidity data
are compared with i
factors suspected to

·RRELATION COEFFICIENTS BETWE
:PONENTS AND DEATH RATES FR
ALL CAUSES OTHER THAN DISEA
› B-28) AND PER CENT OF DEAT
LL-DEFINED, AND UNKNOWN CAUS
(B-45)*
Males 55-59 Years)

All Causes	All Causes Except B-25, B-26, B-27, B-28	Deaths in Category B-45 As Percent of All Deaths
.189	−0.530	−0.549
.340	−0.674	−0.453
.169	−0.466	−0.592
.339	0.296	0.461
067	−0.398	−0.487
099	−0.505	−0.572
275	0.452	0.200
294	0.172	−0.469
372	−0.657	−0.375
178	−0.481	−0.592
277	−0.090	0.632
235	−0.086	−0.275
106	−0.405	−0.479
87	0.521	0.442
96	0.671	0.352

CRITICAL VALUES OF r

N	$\alpha = 0.05$	$\alpha = 0.02$
21	±0.438	±0.521
22	±0.428	±0.508

al food balance data by F.A.O.

France.

nponents which according
coefficients appeared to be
heart disease are indices
attributes of the various
may be that the amount
ilable for consumption is
's development, industri-
lically, and no doubt in
Some of these environ-
ves may be related to
or in fact, there may be
heart disease and animal
neither association has
lid). It may also be
ore abundant diet are
nd diagnostic acumen is
ossible that in some of
ess protein and fat are
ntage of the deaths from

arteriosclerotic and degenerative heart disease are recorded under the noncardiac groupings.

This possibility is supported to some extent by the fact that Table IV shows almost no association between dietary fat or protein and mortality from "all causes of death." Since diseases of the heart form such a large proportion of deaths in the age group fifty-five to fifty-nine, if there is any association between dietary components and heart disease mortality, it should show up also in their correlation with all causes of death. That there is no such correlation again suggests that the association with heart disease is not specific. Moreover, as Table IV shows, there are appreciable negative correlation coefficients between dietary components and death rates from B-45 ("senility, ill-defined and unknown causes"). The latter category may be considered a rough index of the accuracy of cause of death certification in the different countries. The negative association with protein and fat is further evidence of the non-"specificity" of the presumed association.

In summary, the evidence from 22 countries for which data are available indicates that the association between the percentage of fat calories available for consumption in the national diets and mortality from arteriosclerotic and degenerative heart disease is not valid; the association is specific neither for dietary fat nor for heart disease mortality. Clearly this tenuous association cannot serve as much support for the hypothesis which implicates fat as an etiologic factor in arteriosclerotic and degenerative heart disease.

Comment

The problem of identifying causal factors in chronic diseases of unknown etiology is exceedingly complex and can seldom be solved by direct or decisive experiment. Many different types of evidence must be brought together, sifted, analyzed, and tested for possible mutual relationships before an etiologic pattern can be perceived. The "indirect method," therefore, can be a valuable tool in this field of research when used properly and in conjunction with other methods.

The basic material for an indirect study usually comprises statistics on groups; often mortality or morbidity data derived from vital statistics are compared with indices of the environmental factors suspected to be relevant. The data are examined for correlations and concurrent variations to determine whether there is an *association* between the variables.

Such an association does not by itself constitute "proof" of a cause-effect relationship. It merely serves either as a guide for further research or as supporting evidence for a hypothesis which has been formulated on the basis of other theoretic or empiric evidence in the sense that it "fits" with the formulated etiologic pattern.

But herein lies the pitfall. Because it is well understood that an "association" is at best only a small part of the etiologic story, there is a tendency to deal with associations uncritically or even superficially. Investigators must remember that evidence which is not inherently sound cannot serve even for partial support. Evidence for an association must be carefully scrutinized and scientifically weighed. As in any other study the primary data must be evaluated for relevance, accuracy, and internal consistency, and the results must be tested for validity. Whether it is used as a guide by the investigator himself or offered as a link in a chain of supporting evidence for a hypothesis, the association itself must be valid, i.e., "specific" for the variables under study. If the association is not specifically related to the variables but rather reflects a relationship with extraneous or irrelevant factors, it is worse than useless.

Therefore, when an investigator reports an association between certain variables, it is essential that he carry out fundamental procedures of testing and evaluation. The worker who reports the results of an indirect study is under no less obligation to observe methodologic precautions than is the one who reports on the results of any other scientific inquiry. He is obliged to report the basis on which the primary data were selected, their limitations, any qualifying conditions or considerations, and the methods used for testing the validity of the results.

In the proposition considered in this paper—the suggested association between fat in the diet and heart disease mortality—the examination of all available basic data and the tests for specificity show that the association lacks validity. Consequently the apparent association in itself cannot serve as supporting evidence for the theory that dietary fat plays a role in heart disease mortality.

On the other hand, the fact that no association can be demonstrated from this analysis is not in itself sufficient proof that there is no relationship between dietary fat and mortality from arteriosclerotic and degenerative heart disease. The very weakness in the data which give rise to an apparent association when none exists may fail to detect the existence of a valid association at other times. At the present time it is possible to conclude only that international statistics on diet and mortality are not sufficiently sensitive to contribute materially to our knowledge concerning their relationship.

Summary and Conclusions

On the basis of an indirect study reported in 1953, it has been assumed that there is a strong association between the percentage of calories available as fat in the national diet and the national death rates from arteriosclerotic and degenerative heart disease.

The purpose of the present paper is twofold: (1) to evaluate the primary data on which the above study was based and to test the "specificity" of the presumed association and (2) in so doing, to indicate the kinds of supplementary investigation which are essential if an indirect study of association is to yield dependable results.

Analysis of all available data shows that:

1. The apparent association is greatly reduced when tested on all countries for which data are available instead of the six countries used by another investigator.

2. The basic data are subject to considerable limitations. This applies both to the components of the diet in the different countries and to mortality, especially to the classification of causes of death.

3. The presumed association is not "specific" for fat in the diet or for diseases of the heart; for example, the association with heart disease mortality is stronger when animal protein is substituted for fat, and a strong negative association is found for both animal protein and fat with mortality from noncardiac diseases.

4. It is concluded that the suggested association between national death rates from heart disease and percentage of fat in the diet available for consumption cannot at the present time be accepted as valid.

5. It is suggested that in indirect studies of association it is the responsibility of the investigator to report the basis on which the primary data were selected, their limitations, any qualifying conditions or considerations, and the method used for testing the validity of the results.

The authors wish to thank Dr. K. K. P. N. Rao, chief, Food Consumption and Planning Section, Nutrition Division, Food and Agriculture Organization of the United Nations, for making available the data on the components of the diet in the different countries and Miss Helen Supplee for assistance in collecting the data and in statistical analysis.

References

1. Keys, A.: J. Mt. Sinai Hosp. 20: 134 (July-Aug.) 1953.
2. Leitner, Z. A.: Lancet 1: 1079 (May 22) 1954.
3. Keys, A., and Anderson, J. T.: Symposium on Atherosclerosis, National Academy of Sciences, National Research Council, 1955, Publication 338.
4. Rao, K. K. P. N., chief, Food Consumption and Planning Section, Nutrition Division, Food and Agriculture Organization, United Nations: Personal communication.
5. World Health Organization: Annual Epidemiological and Vital Statistics Report, Geneva, 1950–1953.
6. *Ibid.*, 1956, No. 10, pp. 9, 537.
7. Olds, E. G.: Ann. Mathematical Statistics 20: 117 (1949).

(*Appendix A and B follow on pages 2353 and 2354*)

TABL

All countries were utilized fo and for which the necessary in period 1951–1953.

Country
1. Australia
2. Austria
3. Canada
4. Ceylon
5. Chile
6. Denmark
7. Finland
8. France
9. German Federated R
10. Ireland
11. Israel
12. Italy
13. Japan
14. Mexico
15. Netherlands
16. New Zealand
17. Norway
18. Portugal
19. Sweden
20. Switzerland
21. United Kingdom
22. United States

* Source of mortality and
† Age adjustment was ma
population.
The heart diseases includ
Revision of the Internationa
other diseases of the heart.

TABLE II.—Co

Country	So
1. Australia	3,
2. Austria	2,
3. Canada	3,
4. Ceylon	1,
5. Chile	2,
6. Denmark	3,
7. Finland	3,
8. France	2,
9. German Federated Republic	2,
10. Ireland	3,
11. Israel	2,
12. Italy	2,
13. Japan	2,
14. Mexico	2,
15. Netherlands	2,
16. New Zealand	3,
17. Norway	3,
18. Portugal	2,
19. Sweden	2,
20. Switzerland	3,
21. United Kingdom	3,
22. United States	3,

* Data on caloric, fat, an
supplied by Dr. K. K. P. N
culture Organization of the
† Not available.

'y to the classification of

l association is not "specific"
or for diseases of the heart;
sociation with heart disease
when animal protein is sub-
a strong negative association
nimal protein and fat with
ardiac diseases.
d that the suggested asso-
onal death rates from heart
e of fat in the diet available
not at the present time be

that in indirect studies of
esponsibility of the investi-
asis on which the primary
eir limitations, any qualify-
derations, and the method
idity of the results.

nk Dr. K. K. P. N. Rao, chief,
ning Section, Nutrition Division,
nization of the United Nations,
a on the components of the diet
d Miss Helen Supplee for assist-
nd in statistical analysis.

nai Hosp. 20: 134 (July-Aug.)
1: 1079 (May 22) 1954.
rson, J. T.: Symposium on
ademy of Sciences, National
ication 338.
, Food Consumption and Plan-
vision, Food and Agriculture
: Personal communication.
tion: Annual Epidemiological
eneva, 1950-1953.
9, 537.
athematical Statistics 20: 117

Appendix A

TABLE I.—MORTALITY IN 22 COUNTRIES FROM HEART DISEASE FOR MALES*

All countries were utilized for which the dietary data were provided by the Food and Agriculture Organization of the United Nations and for which the necessary information to calculate age-specific rates was available. Wherever possible, rates are for the three-year period 1951–1953.

Country	Rates per 100,000 by Age: 40 to 44	45 to 49	50 to 54	55 to 59	60 to 64	65 to 69	70 to 74	Age Adjusted† (40 to 74)
1. Australia	79.4	176.4	356.4	649.2	1,062.8	1,651.7	2,469.6	654.1
2. Austria	43.3	89.4	207.3	355.9	604.4	1,017.6	1,529.0	384.3
3. Canada	106.0	208.3	383.2	631.6	991.6	1,389.0	2,059.4	606.8
4. Ceylon	41.3	67.9	133.2	173.7	271.0	398.4	476.7	171.5
5. Chile	169.7	279.2	374.4	603.0	1,001.2	1,632.3	2,281.2	667.8
6. Denmark	36.1	88.6	188.9	330.6	581.6	913.5	1,661.9	372.6
7. Finland	131.6	256.9	437.9	757.1	1,168.1	1,863.2	2,654.7	755.0
8. France	39.9	79.9	165.1	282.1	487.9	819.8	1,345.1	319.1
9. German Federated Republic	47.8	90.8	183.3	315.6	515.1	809.3	1,294.5	329.5
10. Ireland	73.0	145.7	288.2	490.6	794.4	1,393.8	2,374.4	550.3
11. Israel	64.6	131.2	283.1	457.5	800.7	1,070.7	1,666.7	461.9
12. Italy	49.6	95.2	175.2	283.4	457.0	757.8	1,355.5	317.3
13. Japan	40.2	64.3	102.6	175.3	281.0	479.2	698.0	189.5
14. Mexico	71.9	108.3	172.2	269.7	319.0	571.2	703.5	243.5
15. Netherlands	32.0	61.3	144.8	242.1	419.9	728.0	1,285.8	284.0
16. New Zealand	69.5	177.3	325.0	525.3	969.4	1,481.0	2,196.6	583.5
17. Norway	27.3	74.0	145.9	256.7	423.2	696.6	1,141.0	275.1
18. Portugal	45.2	77.9	137.7	237.7	413.8	713.8	1,244.3	283.3
19. Sweden	33.5	89.9	167.9	331.0	572.3	973.6	1,610.2	369.7
20. Switzerland	51.4	110.0	206.4	331.3	585.7	1,051.9	1,769.1	404.0
21. United Kingdom	54.1	124.4	257.0	454.8	797.3	1,338.0	2,216.3	516.0
22. United States	136.6	267.1	478.5	774.2	1,152.8	1,695.4	2,353.7	729.2

* Source of mortality and population data: WHO, *Annual Epidemiological and Vital Statistics*, 1951–1953.[5]
† Age adjustment was made by the direct method, using the sum of the populations of the 22 countries as the standard population.
The heart diseases included in this group are those corresponding to the abbreviated list numbers B-26 and B-27 (1948 Revision of the International Lists of Causes of Death); B-26 is arteriosclerotic and degenerative heart disease, and B-27 is other diseases of the heart.

TABLE II.—COMPONENTS OF NATIONAL AVERAGE FOOD SUPPLIES AVAILABLE FOR HUMAN CONSUMPTION*

Country	Calories per Person per Day: All Sources	Protein: Total	Protein: Animal	Fat: Total	Fat: Animal	Carbohydrate	Calories as Per Cent of Total: Protein: Total	Protein: Animal	Fat: Total	Fat: Animal	Carbohydrate
1. Australia	3,290	380	252	1,098	954	1,812	12	8	33	29	55
2. Austria	2,770	332	160	855	684	1,583	12	6	31	25	57
3. Canada	3,110	380	236	1,170	909	1,560	12	8	38	29	50
4. Ceylon	1,790	172	44	306	54	1,312	10	2	17	3	73
5. Chile	2,490	308	108	495	315	1,687	12	4	20	13	68
6. Denmark	3,240	364	200	1,269	936	1,607	11	6	39	29	50
7. Finland	3,100	384	204	927	774	1,789	12	7	30	25	58
8. France	2,790	380	188	810	...†	1,600	14	7	29	...†	57
9. German Federated Republic	2,900	304	160	1,008	639	1,588	10	6	35	22	55
10. Ireland	3,500	380	188	1,080	945	2,040	11	5	31	27	58
11. Israel	2,670	352	104	612	171	1,706	13	4	23	6	64
12. Italy	2,600	320	88	540	261	1,740	12	3	21	10	67
13. Japan	2,010	232	52	162	45	1,616	12	3	8	2	80
14. Mexico	2,270	260	60	513	207	1,497	11	3	23	9	66
15. Netherlands	2,860	324	168	1,044	468	1,492	11	6	37	16	52
16. New Zealand	3,340	408	276	1,341	1,206	1,591	12	8	40	36	48
17. Norway	3,120	360	200	1,179	891	1,581	12	6	38	29	51
18. Portugal	2,440	268	88	603	234	1,569	11	4	25	10	64
19. Sweden	2,980	344	220	1,152	990	1,484	12	7	39	33	50
20. Switzerland	3,080	368	204	1,017	702	1,695	12	7	33	23	55
21. United Kingdom	3,140	340	184	1,188	954	1,612	11	6	38	30	51
22. United States	3,090	364	252	1,215	855	1,511	12	8	39	28	49

* Data on caloric, fat, and protein content of national average food supplies available for human consumption (1951–1954) supplied by Dr. K. K. P. N. Rao, Chief, Food Consumption and Management Section, Nutrition Division, Food and Agriculture Organization of the United Nations.
† Not available.

*Appendix B**

Calorie, Protein, and Fat Consumption Levels in Selected Countries

NATIONAL AVERAGE FOOD SUPPLIES AVAILABLE FOR HUMAN CONSUMPTION (FOOD BALANCE SHEETS) PREPARED BY F.A.O.

National food balance sheets are used for estimating average food supplies available for human consumption as a measure of per caput consumption levels of different foods. The quantities of food produced in a country adjusted by imports, exports, and changes in stocks are equal to the gross food supplies available in that country. If from this amount are deducted quantities representing all non-food uses, including wastage up to the retail stage, the remainder represents the total food supplies available for human consumption. The per caput supplies are obtained by dividing these figures by the number of population and the supplies are then expressed in terms of calories, protein, and fats.

Since estimates of national average levels of food consumption are thus obtained through food balance sheets as residual quantities, it means that their validity depends on the reliability of the national statistics on production, marketing, and utilization. Hence, the accuracy of data on different countries is very variable and depends largely on the development of the countries with respect to their statistical services. In most of the less developed countries, food balance sheets cannot be relied on to produce more than rough indications of the average consumption levels of the main basic foods. Data on animal products are particularly weak in this respect.

Apart from the inaccuracy and inadequacy of the statistical data on which food balance sheets are based for many countries, their main limitation is that they indicate only national average food supplies available for human consumption and do not represent actual consumption levels. Moreover, these national average figures conceal vital differences in consumption of different sections of the population within a country, i.e., by regions, income levels, occupations, etc. Such detailed information, which is often essential for nutritional purposes, can be obtained only through specific surveys on representative samples of various sections of the population.

* Excerpt from Working Paper No. 6, WHO Study Group on Atherosclerosis and Ischemic Heart Disease, Geneva, November 7 to 11, 1955.

Doctor Contributions to Medical Schools

The American Medical Education Foundation reports that physicians gave well over three million dollars to medical education in 1956.

The AMEF just released data giving a breakdown of physician contributions to medical education last year. For the first time, this also includes information on contributions made through alumni campaigns. The report showed:

In 1956, 84,657 doctors gave a total of $3,320,152.14 to the country's 83 medical schools. This total included $1,072,727 given through the AMEF by 39,892 doctors, and $2,247,425 given directly to the medical schools by 44,765 doctors.

The AMEF's million-plus contribution is to be used at the discretion of the schools. The new information shows that most of the contributions made through alumni campaigns are also "unmarked," that is, they may be allocated as the deans of the individual schools see fit.—*Secretary's Letter, April 25, 1957*

C. STUART WE

(From the Departments

INDICATIONS for sp
over the years as w
some of the pathologi
and as we have increa
surgery within the ab
indications for splenec
group, as can easily b
of the surgical pathol
tory. In a six-year p
sive) 186 splenectom
Albany Hospital. T
ruptured spleen, 58 f
100 so-called "inciden
the largest single grou
A discussion of in
can be conveniently
medical indications

Surgical Indicati

There is little new
cations for splenecto
ment is true in regar
but the ever-increasi
performed in conjunc
in the upper abdom
First of all, if the s
section, it should be
sure of hemostasis.
uncertain and dange
tomies in this group
adequate resection
instances it makes t
easier in cases of be
indication for splen
gastrectomy is in su
to attain as near a
anatomic situation
body and fundus
requires distal panc
in order to remove
course of the splen
additional excision
ness of total gastr