PRINCIPLES OF MEDICAL STATISTICS

II.—SELECTION

In medical statistics we are nearly always working with samples drawn from large populations. For instance if we compare the mortality of diabetics treated with insulin with diabetics not treated with insulin, we do not possess information regarding all such persons (whom we may term the "universe" of diabetics) but only for a sample of each type which has come under our own observation or under the observation of others. If we wish to argue from this sample to the universe from which it was drawn, to deduce that what is true of our sample of patients is therefore true of the general run of patients, then we must consider very carefully whether in fact our sample is fully representative of all patients, and not in any way biased or "selected." It is important to be clear on the meaning the statistician attaches to the word selected. By a selected sample he denotes a sample which is not representative of the universe from which it is drawn. The selection may have been deliberate, in which case the form of selection is known and the lack of comparability between the sample and the universe is usually perfectly clear. For instance, if the treatment of respiratory tuberculosis by means of artificial pneumothorax is confined in a sanatorium to patients with signs of disease in one lung only, then it is obvious that these patients are *not* a representative sample of all patients with phthisis, but are selected on the criterion of one lung only affected. To compare their mortality experience with that of all patients is, therefore, a very doubtful procedure, for we are clearly not comparing like with like in all respects except artificial pneumothorax treatment. Even without that treatment the death-rate of patients with only one lung affected may differ materially from the death-rate of the general run of patients.

More often, however, the "selection" is not deliberate but is quite unforeseen or is unrealised. To say, therefore, that Mr. So-and-So's figures relate to a selected sample of patients is not an aspersion on Mr. So-and-So's scientific honesty; the statement implies only that owing to the method of collection of the figures, or to the limited field in which Mr. So-and-So was able to operate, it is quite impossible for his sample to be representative. It may be that with care that selection might have been avoided; often it is unavoidable. Its possible presence cannot be too carefully remembered or taken into account in interpreting statistics. As however it is frequently overlooked a series of examples will not be amiss.

Examples of Selection

I

As a simple illustration I have taken from the *Times* the frequency with which male and female births are recorded in the birth column. In 1935 the number of male births was 3304 and female births 3034, so that the sex ratio is 1089 males for each 1000 females. According to the Registrar-General's figures for England and Wales the sex ratio of births in the country as a whole rarely exceeds 1050. It is clear that from the point of view of sex ratio the births recorded in the *Times* are not representative of the births in the country as a whole. It is possible that first births are more frequently recorded in those columns than births of a later order, and that such births have a different sex ratio; or that proud parents are more likely to record their heirs than their daughters; or that the sex ratio differs between social classes. With such a sample of births, if that was all that was available, one could not generalise about the universe with any security.

II

Hospital statistics can very rarely be regarded as unselected. The patients are frequently drawn from particular areas and largely from particular social classes. Still more important, in many diseases only those patients who are seriously ill are likely to be taken to hospital. It is obvious that we cannot determine with any approach to accuracy the fatality-rate of any disease, say, measles at ages 0–5 years, if our statistics are based only upon the seriously ill—perhaps only upon patients in whom a secondary pneumonia has developed—and ignore the mass of children whose symptoms are so slight that they can safely be treated in their own homes. Of *all* children with measles, those in hospital will often form only a small and stringently selected group; our deductions from such a group are correspondingly limited, especially with regard to such factors as the incidence of complications and the rate of fatality or recovery. It is not too much to say that there is no disease in which a hospital population must not initially be regarded with suspicion if it is desired to argue from the sample to the universe of all patients. No such argument should be attempted without a preliminary and rigorous examination of the possible ways in which selection may have occurred.

The same difficulty arises with secular comparisons—e.g., when we wish to see whether the fatality from some disease has changed from one year to another. In each year the fatality-rate is measured upon the patients admitted to hospital, and in each year those patients are a sample of all patients with the disease in question. It must be considered whether that sample has changed in type. In both years the sample may be a selected sample but the selection may not be identical. The kind of patient admitted may have changed. For example, in a group of American hospitals it has been reported that the fatality-rate from appendicitis declined from 6 per cent. in 1928 to 3·5 per cent. in 1932. Is that a "real" decline or has there been a concurrent change in the types of patients admitted? Examination of the basic figures shows that in these hospitals in 1928 some 2500 patients were operated upon while in 1932 the number had risen to 3500, an increase of 40 per cent. It is impossible to believe that an increase of 40 per cent in five years is a real increase in the incidence of appendicitis. It is more likely that the desire for admission to these hospitals or the criteria of admission have changed, that some patients who were admitted in 1932 would not have entered them in 1928. It is possible, therefore, that the type of entry has changed as well as the volume—perhaps that milder cases were admitted and operated upon in 1932 which were not present in the 1928 series. In the absence of positive evidence on that point the change in fatality cannot be accepted at its face value or as satisfactory evidence of the effect of a change in some other factor—e.g., the benefit of earlier admission to hospital in 1932 than in 1928. The following two questions must always be considered. Has there been a change in the population from which the samples are drawn at two dates—i.e., a change

relevant to the question at issue? At each date was there an equal probability that a particular type of patient would be included in the sample?

III

In measurements of the value of some form of treatment, statistics of the following type are frequently given :—

Day of disease upon which treatment was first given.	Fatality-rate per cent. of treated patients.	Day of disease upon which treatment was first given.	Fatality-rate per cent. of treated patients.
1	1·3	4	9·3
2	3·6	5	12·8
3	7·5	6 or later	16·4

It is possible that the level of this fatality-rate at the different stages is seriously influenced by selection. Let us suppose, as is often the case, that the treatment is given to patients brought to hospital and that *all* patients do not necessarily go to hospital. Then on the first day of disease a variety of patients will be taken to hospital, in some of whom, in the absence of the special treatment, the disease is destined to run a mild course, in others a severe course. The presence of a proportion of mild cases will ensure a relatively low fatality-rate, even if the special treatment has no specific effect. But as time passes this proportion of mild cases in the hospital sample is likely to decline. By the time, say, the fourth day of disease is reached, a number of patients who were not seriously ill will have recovered or be on the way to recovery. Their removal to hospital is unnecessary. On the other hand, patients who have made a turn for the worse or whose condition has become serious are likely to be taken to hospital for immediate treatment. Thus on the later days of disease the sample removed to hospital for treatment is likely to contain an increasing proportion of persons seriously ill, it obviously being unnecessary to transfer those who are making an uninterrupted recovery. In other words the patients removed to hospital on the fourth day of disease are not a random sample of all patients who have reached that day of the disease but consist rather, perhaps mainly, of patients still seriously ill. Such a group will certainly have a relatively high fatality-rate.

Another example of this statistical difficulty may be taken from some fatality-rates recorded for appendicitis. It has been reported that in a group of cases 2 per cent. died of those admitted to hospital within 24 hours of the onset of symptoms compared with 10 per cent. of those whose admission was delayed till after 72 hours. But it is likely that the group of patients admitted early is composed of a proportion of the seriously ill and a proportion that would do well whether admitted to hospital or not. On the other hand those who are admitted after a delay of three days from onset are likely to be patients whose condition is serious, and clearly those whose condition has become quiescent are unlikely to be taken to hospital at that point of time. If such a sequence of events occurs, it is clear that the group of patients admitted early is not in pari materia with the group of patients admitted late. Selection may not be the whole explanation of the difference between the fatality-rates, indeed it is not likely to be for there are excellent reasons for the early treatment of appendicitis. But it is a possible factor with statistics such as these which makes it very difficult to measure accurately the *magnitude* of the advantage.

IV

A sample which is composed of volunteers or self-selected individuals is not likely to be a random sample of the universe from which it is drawn. If, for example, the treatment of colds by vaccine is offered to a group of persons, the volunteers are likely to belong mainly to that section of the group which suffers most severely from colds and hopes for some advantage from the treatment. They are in that event a select group, not comparable with the remainder of the population from which they were drawn. In such cases the question must always arise: is the act of volunteering correlated with any factor which may have an influence upon the final results of the experiment?

Self-selection

An interesitng example of what may be termed self-selection is worth quoting. In 1931 the Industrial Pulmonary Diseases Committee of the Medical Research Council wished to make an inquiry into whether the working health and capacity of coal-miners are impaired by the inhalation of anthracite dust. To begin with, a study was made of the size and age constitution of the working population at the South Wales anthracite collieries. Underground workers had to be excluded from the investigation, since such workers as well as being exposed to various concentrations of anthracite dust in the atmosphere may in addition be exposed at times to silica dust. If the health of these workers was found to be impaired it would be impossible to implicate anthracite dust as the responsible agent. The impairment might equally well be due to exposure to stone dust containing silica, which, it is well known, can produce serious damage to health. In addition it was considered necessary to exclude surface workers who had at any time worked underground, since the effects of exposure to silica dust will not necessarily be immediately apparent and also because impaired health may have been the reason for transference from underground to surface work. This was, in fact, known to be the reason in numerous cases, so that such workers would be a highly select group. Attention was therefore turned to workers who were exposed to anthracite dust on the surface and had *always* worked on the surface. Such workers, it was found, are employed on a relatively light task. Not only was there a tendency to draft to it operatives who had previously worked underground and had for one reason or another become partially incapacitated, but in addition it was clear that a large number of boys were initially employed upon this work but rapidly moved away to other work. In the main these boys were drafted underground where the physical labour was heavier but the rate of pay superior.

The inevitable inference is that the healthy and strong individuals will transfer to underground work while those who remain on the surface are likely to be of under-average physique and health. In other words there has been a form of self-selection. If the examination of such surface workers showed that they included a high proportion with impaired health or that they suffered an unduly high rate of sickness in comparison with some standard, this result could not with security be ascribed to the effects of dust inhalation. It might be considerably influenced by the fact that these surface workers were, through the operation of selection, initially less healthy than a random sample of all surface workers. This investigation also contains an example of selection through

volunteers being accepted for examination. For the reasons outlined above no inquiry was made at the collieries and the field of study was transferred to dock-workers exposed to anthracite dust. Of 250 such workers it was arranged to examine, clinically and radiologically, a sample of 40 operatives—namely, 15 workers employed for only 3–4 years and 25 older workers with 15–40 years' service. These two groups were selected at random from the complete list of operatives, to ensure, as far as possible with such small numbers, a representative sample. At the examination eleven of these men were absent and to make good the deficiency in numbers volunteers were secured in place of the absentees.

The results of the examinations suggested that there was a readiness to volunteer on the part of individuals who, on account of some known or suspected disability, desired to be medically examined. Such substitutions may, therefore, result in the sample ceasing to be random and representative of the population from which it was drawn.

Questionnaires

Inquiries carried out by means of questionnaires are par excellence those in which selection must be suspected. In all such inquiries replies to the questions put—even to the simplest question—are received from only a proportion of the individuals to whom the form is sent. There can never be the slightest certainty that the individuals who choose to reply are a representative sample of all the individuals approached; indeed very often it is extremely unlikely that they are representative. For example one may take the inquiry made by the Editor of THE LANCET into the Present-day Openings of Medical Practice (*Lancet*, 1935, **2**, 512). To measure the success with which recent graduates had been attended in their profession a questionnaire of three relatively simple questions was addressed to the 1490 men and women who in 1930 registered their names with the General Medical Council, viz., (1) What branch of medicine have you taken up? (2) What led you to this choice? and (3) What was your approximate income from professional work last year? To overcome objections to providing such personal information no clue to the identity of the correspondent was required. Of the individuals approached 44 per cent. replied. Are these persons a representative sample of the 1490? It is possible, as is clearly pointed out in the report, that there might be a tendency for those who have been successful in their profession to be more eager to register their success than for those who have failed to register their failure. Alternatively the latter might under the veil of secrecy be glad of the opportunity of stating frankly the drawbacks of the profession. Those who have turned to other professions might tend not to reply under the impression that the inquiry cannot concern them. Successful and busy individuals might be unwilling to give time to the inquiry. It is impossible to determine whether any such factors are operative in the determination to answer or not to answer. The difficulty is inherent in all inquiries carried out by this method.

House Sampling

An interesting example of selection in taking a random sample of houses is suggested in the Ministry of Health's report on the influenza pandemic of 1918 (Reports on Public Health and Medical Subjects, No. 4). To obtain facts as to the incidence and fatality from influenza in 1918–19 a house-to-house inquiry was undertaken in five areas of Leicester, information being obtained *so far as possible* at every fifth house. Houses which were found closed at the time of visit had to be ignored in this census. But houses in which there are young children are rarely found closed and this would tend to affect the age-distribution of the population recorded in the sample. Compared with the population from which it was drawn the sample would be likely to contain an undue proportion of young children and a deficit in the number of adults. Any substitution of another house for the original randomly chosen one would be likely to add still further bias to the sample, and such substitutions are to be avoided in sampling inquiries.

It will be noted that a selection of this type would be difficult to foresee. It is here that the statistician has some advantage, for his experience of such inquiries makes him familiar with the methods that are likely to ensure a random sample and those that are likely to lead to one that is unrepresentative of the population from which it is taken. Workers who are unfamiliar with sampling inquiries but wish to embark upon one may, therefore, find his advice of assistance.

Summary

In statistics we are working, nearly always, with relatively small samples drawn from large populations; to be in a position to generalise the sample must be representative of the population to which it belongs. In taking samples "selection" may occur through the operation of various factors. A selected sample is one which is not representative of the universe, whether the bias be due to deliberate choice or unconscious selection of the members incorporated in the sample. In generalising from a sample, or in making comparisons between one sample and another, the possible presence of selection must always be very closely considered.

A. B. H.

INFECTIOUS DISEASE

IN ENGLAND AND WALES DURING THE WEEK ENDED DEC. 26TH, 1936

Notifications.—The following cases of infectious disease were notified during the week: Small-pox, 0; scarlet fever, 1636; diphtheria, 1013; enteric fever, 21; pneumonia (primary or influenzal), 832; puerperal fever, 22; puerperal pyrexia, 72; cerebro-spinal fever, 16; acute poliomyelitis, 3; encephalitis lethargica, 4; dysentery, 16; ophthalmia neonatorum, 40. No case of cholera, plague, or typhus fever was notified during the week.

The number of cases in the Infectious Hospitals of the London County Council on Jan. 1st, 1937, was 3451, which included: Scarlet fever, 999; diphtheria, 1066; measles, 10; whooping-cough, 427; puerperal fever, 14 mothers (plus 6 babies); encephalitis lethargica, 285; poliomyelitis, 4. At St. Margaret's Hospital there were 12 babies (plus 5 mothers) with ophthalmia neonatorum.

Deaths.—In 122 great towns, including London, there was no death from small-pox, 2 (0) from enteric fever, 1 (0) from measles, 2 (0) from scarlet fever, 11 (4) from whooping-cough, 40 (7) from diphtheria, 33 (8) from diarrhœa and enteritis under two years, and 97 (25) from influenza. The figures in parentheses are those for London itself.

Sheffield and Stockton-on-Tees each reported a death from enteric fever; Thurrock 2 deaths from whooping-cough. Fatal cases of diphtheria were reported from 18 great towns, Liverpool 6, Thurrock 4, Bradford and Manchester each 3.

The number of stillbirths notified during the week was 220 (corresponding to a rate of 49 per 1000 total births), including 33 in London.