# PRINCIPLES OF MEDICAL STATISTICS

## XIII—FURTHER FALLACIES AND DIFFICULTIES

### Proportional Rates

IT is not uncommon for some confusion to be shown over the difference between absolute and relative figures. For instance, take the following figures for England and Wales in 1932.

| Age-group (years). | Estimated population in age-group. | Number of deaths registered. | | Death-rates per million. | |
|---|---|---|---|---|---|
| | | All causes. | Tuberculosis (all forms). | All causes. | Tuberculosis (all forms). |
| 5– 9 | 3,181,900 | 6601 | 774 | 2075 | 243 |
| 10–14 | 3,349,100 | 4643 | 827 | 1386 | 247 |

From the estimated populations and the numbers of deaths registered the death-rates experienced in the two age-groups can be calculated. At ages 5–9 the death-rate from tuberculosis was 243 per million persons and at ages 10–14 it was 247, a negligible difference. In other words the risk of dying from tuberculosis did not vary materially in 1932 between these two age-groups. Supposing, however, we had no means of estimating the populations exposed to risk, and the sole information at our disposal consisted of the deaths registered from tuberculosis and from all causes. In that event the only rate that it is possible to calculate is a proportional rate—i.e., the deaths from tuberculosis as a percentage of the deaths from all causes. At ages 5–9 this rate is $774 \times 100 \div 6601 = 11{\cdot}7$ per cent. and at ages 10–14 it is $827 \times 100 \div 4643 = 17{\cdot}8$ per cent. The *proportional death-rate* from tuberculosis is 52 per cent. higher at ages 10–14 than at ages 5–9. It is necessary to be quite clear on the meaning of this proportional rate. The information it gives is that *in relation to other causes of death* mortality from tuberculosis is more frequent at ages 10–14 than at ages 5–9. This is certainly an important aspect of the figures; it clearly shows that any action which reduces deaths from tuberculosis (without changing the death-rate from other causes) will have more effect on the total death-rate at ages 10–14 than at ages 5–9, because more of the total deaths are attributable to tuberculosis at the higher ages. But obviously we cannot deduce from the proportional rates that the *absolute* risk of dying from tuberculosis differs in the same ratios; we have, in fact, seen that the absolute risks are equal. The proportional rate is dependent upon the level of two factors—viz., the deaths from tuberculosis and the deaths from all causes—which form the numerator and denominator of the ratio. Differences in *either* factor will influence the ratio. In the case cited the numerators (deaths from tuberculosis) are nearly equal, 774 at ages 5–9 and 827 at ages 10–14, but the denominators (deaths from all causes) differ appreciably, 6601 at 5–9 and 4643 at 10–14. The proportion of deaths due to tuberculosis at ages 10–14 is higher than the proportion at ages 5–9, not because of a rise in the incidence of tuberculosis but because there is a fall in the incidence of other causes of death. While other causes of death—e.g., diphtheria, scarlet fever, measles—are declining, tuberculosis remains nearly steady. Relatively, therefore, its importance increases, absolutely it is unchanged.

### Changes in Proportional Rates

To take another example, in 1917 the deaths at age 0–1 year formed 12·9 per cent. of all the deaths registered in England and Wales, whereas in 1918 the proportion was 10·5 per cent. It does *not* follow that the infant mortality-rate of 1918 was lower than that of 1917; in fact it was very slightly higher, 97 per 1000 live births compared with 96 in 1917. The lower *proportion* in 1918 is due to the great increase in deaths *at ages over 1* as a result of the influenza pandemic. The numerator of the ratio did not change appreciably (64,483 deaths under 1 in 1917 and 64,386 deaths in 1918); the denominator did change appreciably (from 498,922 deaths at all ages in 1917 to 611,861 deaths in 1918).

Similarly an absence of change in a proportional rate does not necessarily denote an absence of change in the absolute rate. If in 1900 there were 5 per cent. of all deaths due to cause X and in 1930 the proportion was still 5 per cent., this is not incompatible with an absolute decline in the mortality due to cause X. If the death-rates due to cause X and to all other causes have *both* been halved the proportion must remain the same.

#### PROPORTIONAL DEATH-RATES IN OCCUPATIONS

Correspondingly a higher proportion of deaths due to cause X amongst, say, coal-miners than amongst bank clerks does not necessarily denote an absolutely higher death-rate from that cause amongst the former. Expressing the numbers of deaths due to each various cause as percentages of the total number due to all causes gives a series of figures that must add up to 100. Differences between any one proportion amongst miners and bank clerks must to some extent influence all the other proportions. For instance, if the proportion of deaths amongst miners which is due to accidents is very much higher than the corresponding proportion amongst bank clerks it follows that other causes of deaths amongst miners must be *proportionately* decreased. The point is clear if we suppose only two causes of death to exist, say accidents and tuberculosis. Amongst populations of clerks and miners of equal size there are, let us suppose, in the former 50 deaths of each kind, while amongst miners there are 100 deaths from accidents and 50 deaths from tuberculosis. The actual mortality-rates from tuberculosis (to deaths divided by the number in the population) are identical for the populations are of equal size; but proportionately the tuberculosis deaths form only one-third of the total amongst miners and one-half amongst clerks. The relative excess of tuberculosis amongst clerks is not due to an absolutely greater risk of this disease but to a lower risk from accidents.

The cardinal rule in the interpretation of proportional rates is to pay equal attention to the numerator and to the denominator of the ratios. Departure from that rule may be illustrated in one or two examples culled from published figures.

#### PROPORTIONAL RATES IN HOSPITAL STATISTICS

In hospital statistics the population at risk—i.e., from which the recorded cases are drawn—is usually not known. Mortality or incidence rates cannot, therefore, be calculated, and one falls back on proportional rates.

For example, cases of pernicious anæmia are expressed as a proportion of all cases of illness admitted to hospital, the hospital being situated in the United States. Of the 47,203 admissions of persons who were born in the United States 291, or 0·62 per cent., were cases of pernicious anæmia; of the 2814 admissions of persons who were born in England 25 were cases of pernicious anæmia or 0·89 per cent.; and of the 7559 admissions of persons who were born in Russia 14 were cases of pernicious anæmia or 0·19 per cent. It is impossible to accept these figures as adequate evidence of racial differences in the liability to pernicious anæmia. The mortality from that disease is correlated both with age and with sex, the death-rate rising with age and falling more heavily on women than on men. There is no doubt that the incidence varies similarly. The numerator of the proportional ratio (the number of cases of pernicious anæmia) is therefore influenced by the sex and age composition of the population from which the cases are drawn, and it is unlikely that the composition is the same in native-born and immigrant populations.

Equally the age and sex compositions of the populations at risk in the area may influence the admission rates for other causes of illness—i.e., the denominators of the ratios—since many causes of illness are correlated with age and sex. In the absence of any knowledge of the constitution of those populations it is impossible to draw reliable conclusions from these different proportions of pernicious anæmia.

PROPORTIONAL RATES IN INFANCY

The infant mortality experienced at different stages of the first year of life was differentiated between urban and rural areas and gave the following figures :—

*Deaths per 1000 Live Births at Different Months of Age*

| — | Under 1 month. | Months. | | | | Under 1 year. |
|---|---|---|---|---|---|---|
| | | 1– | 2– | 3– | 6–12 | |
| Urban .. | 29·67 | 12·06 | 8·73 | 22·77 | 22·14 | 95·37 |
| Rural .. | 23·77 | 7·44 | 5·18 | 10·98 | 11·29 | 58·66 |

Expressing these figures as proportions it is concluded that "it is significant that while the proportion of infants succumbing in the first month after birth was 31 per cent. in urban areas, it rose to 41 per cent. in rural areas, which points to the lack of institutional facilities in connexion with confinements in such districts and to the difficulty often experienced of summoning skilled medical assistance in time." Clearly that conclusion does not necessarily follow from the proportions. The *absolute* rate at under 1 month is *lower* in the rural areas than in the urban areas (23·77 to 29·67)—though possibly it might be lower still given better institutional facilities. But it will remain *proportionately* higher just as long as the urban environment shows a relatively excessive death-rate at ages 1–12 months. If 69 per cent. of the deaths are at ages 1–12 months in the urban areas and only 59 per cent. in the rural areas, then whatever the absolute rates at age 0–1 month the proportions *must* be 31 in the urban and 41 in the rural areas at that age—i.e., a *relatively* unfavourable state of affairs in the rural areas. Given the best possible institutional facilities that state of affairs will prevail so long as the urban areas show higher rates during the remainder of the first year.

From the absolute rates in fact two arguments may be advanced. The smaller advantage in the rural areas at 0–1 month compared with later ages may be the result of inferior institutional facilities there, or it may be due to the fact that the urban environment does not exert its maximum unfavourable effects until after the age of one month.

STATISTICS OF POST-MORTEMS

Attention has been paid in recent years to the question as to whether cancer of the lung has increased. The recorded mortality statistics have been regarded as unsatisfactory evidence, for the rise apparent in those figures may be due to more accurate certification of death now than in the past. The number of post-mortems at which this form of malignant disease is found is held to be a more certain and less changeable basis. This number can be taken as the numerator of the rate. But in the absence of a knowledge of the population from which the hospital draws its inmates, an incidence rate cannot be calculated. In its place a proportional rate is used, taking, for instance, as the denominator either the total number of post-mortems carried out, or the total of all cases of disease admitted to the hospital. Thus we have the cases of cancer of the lung observed at post-mortem as a proportion of all post-mortems or of all cases admitted. Calculating this ratio year by year shows its secular trend. In interpreting this trend we have to consider possible changes both in the numerator and the denominator. The numerator, cases of cancer of the lung found at post-mortems, may be influenced by an increasing interest in the condition ; persons in whom it is now suspected may be more frequently submitted to post-mortem than in past years. If the denominator were unchanged this would give an apparently rising incidence. But it is equally important to inquire whether the denominator has changed—i.e., all post-mortems or all cases admitted. If the criteria upon which post-mortem examination or admission to hospital are based have changed with time this will also inevitably influence the trend of the ratio. An addition to the denominator of cases or post-mortems that would not have been present in earlier years must result in a declining ratio, even though the actual incidence rate of cancer of the lung on the population exposed to risk has not changed.

For example, suppose the incidence of cancer of the lung in the unknown population is 5 cases per annum, all of which come to the hospital, and are recognised, and this number is unchanged between 1900 and 1930. If in 1900 there were 500 post-mortems carried out then the proportional rate is 5 in 500 ; if in 1930 a laudable attempt to determine more accurately the causes of death of patients led to an increase of the post-mortem examination to 600, the ratio falls to 5 in 600, without any change in the real incidence.

It may, perhaps, be reiterated that the denominator needs just as much consideration as the numerator.

## The Problem of Attributes

In medical statistics subdivisions of the cases treated have often to be made in order to ensure that like is being compared with like. It is obviously idle to compare the fate of two groups of patients with cancer of the cervix uteri treated by different methods, if one group contains 50 per cent. of persons whose disease was far advanced when treatment was begun and the other had but 10 per cent. in the same category. To avoid this fallacy an attempt is made to classify the patients according to the stage, or severity, of the disease when first seen. Such a division, it is recognised, may be influenced by the different interpretation of what constitutes degrees of severity by the observers, and such statistics must always be regarded with that difficulty in mind.

For instance, one hospital classifies two-thirds of its patients to stages of disease 1 and 2 and one-third to stages 3 and 4; another hospital has only one-quarter of the patients at stages 1 and 2 and three-quarters at stages 3 and 4. Do patients really present themselves at those hospitals in such widely differing proportions or is the difference due to differing standards of classification? The answer to that question is unknown, but in comparing the fatality-rates following treatment the problem must not be disregarded. It must be remembered that different standards of classification may favourably influence one hospital as compared with another at *every* stage of the disease. This may be shown diagrammatically by supposing the patients are set out in a line in order of severity, from mildest to most severe.

| | | | | |
|---|---|---|---|---|
| Hospital A .. | Stage 1 | Stage 2 | Stage 3 | Stage 4 |
| Hospital B .. | Stage 1 | Stage 2 | Stage 3 | Stage 4 |

Hospital B includes in stage 1 a proportion of patients which Hospital A relegates to stage 2; the latter hospital, therefore, has only the very best of the patients in stage 1 and will in consequence have a more favourable result than Hospital B at that stage. In the stage 2 group Hospital A includes favourable cases which B has called stage 1, while B includes less favourable cases which A classifies to stage 3. Hospital A will again show a better result. The same difference is apparent in the subsequent stages, Hospital B having only the extreme cases in stage 4 and A including some which B would relegate to stage 3. At each stage Hospital A will, therefore, compare favourably with Hospital B, though in the total there may be no difference whatever.

In any series of statistics based upon division by attributes by different observers this possible fallacy must be considered with care. The difficulty occurs acutely in the assessment of malnutrition amongst school-children where different observers use different criteria of determination and place a different meaning on such terms as "good," "fair," "poor," &c. The result has led in the past to an incidence that differs unbelievably from one place to another, but so far no objective measure has been devised which will satisfactorily replace the subjective assessment. In the same way differing interpretations by different persons made the statistics of the blind and dumb collected at the census enumerations in the past of little value, and led to the questions being abandoned. Faced with such statistics the reader must always ask himself: Could these classifications have been made by means of objective measurements, which, given equal skill in making those measurements, would not vary appreciably from observer to observer? If not how far may the presence of subjective influences affect the results and in what way?

A. B. H.

# SPECIAL ARTICLES

## SILICOSIS AT HOME AND ABROAD

A PLEA for the extension of legislation in this country in respect of silicosis was made in a communication to the Medico-Legal Society reported in our columns on March 6th. In this connexion it is of interest to review the progress in the gold-mining industry in relation to the legislation for compensation for silicosis.

### History of Legislation on the Rand

Gold was "proved" on the Rand in 1885, and mining started there in 1886. At first the mine workings were shallow, and as the rock was relatively soft, friable, and damp, the amount of dust was not excessive; but by the year 1892 extensive deep-level mining was being done with an increasing use of drilling machines. In 1899 100 mines were working, employing 12,000 whites and 100,000 coloured persons. Before the Boer War very few people suspected that the mine dust was causing any injury to health. The first warning was given by the Government mining engineer of the Transvaal Mines Department in 1902, when in his report he introduced the subject of the health of the workmen and mentioned miners' phthisis for the first time. He stated that of 1377 machine men employed before the Boer War, 225 were known to have died between October, 1899, and January, 1902. The attention of the Government was arrested by these facts, and in December, 1902, the first Transvaal Miners' Phthisis Commission was appointed by Lord Milner, to inquire into the incidence of "the disease commonly known as miners' phthisis."

Regulations were made for improving the ventilation of the mines, and these were supplemented from time to time by notices. In 1907 the Mining Regulations Commission was appointed to revise and suggest amendments to the Regulations, and a new code was issued in 1911. In the latter year compensation for miners' phthisis was introduced for the first time, in the Miners' Phthisis Allowances Act, to provide temporary relief to necessitous cases pending more extensive legislation. Payment of compensation was made by the last employers who, however, could recover half the amount from the Government and a part of the remainder from a previous employer. In 1912 an Act was passed which laid down a financial basis on a modified insurance method. The benefits were financed by two funds—the Miners' Phthisis Compensation Fund and the Miners' Phthisis Insurance Fund. The former provided benefits for the first 2 years from a fund which consisted mainly of quarterly levies on the "scheduled" mines, and a contribution from the Government to make up any deficit, up to a sum of £100,000. ("Scheduled" mines were those wherein the mineral dust produced by mining operations was, in the opinion of the Minister, of such a nature as to cause silicosis.) The insurance fund was for cases applying for benefit after August 1st, 1914. It was financed by contributions from the mines and from the miners, based on wages; the mines contribution varying from 2½ per cent. to 5 per cent., the miners' contribution being 2½ per cent. The main reason for requiring workers to contribute, apparently, was to secure their coöperation in preventive measures. An amending Act was passed in 1914, extending the provisions for compensation to mine crusher-house workers as well as miners and, in certain circumstances, to native labourers.

An important Act passed in 1916 gave effect to the recommendations of a Select Committee appointed in 1915. "Miners' phthisis" and "silicosis" were now regarded as synonymous terms, and a central medical bureau was established for dealing with all medical questions. The Act prescribed periodical medical examinations, at intervals of six months, of white miners employed underground, and suspension from work in the mines of any person found to have tuberculosis of the lungs.

#### A LANDMARK

The Miners' Phthisis Act of 1919 repealed all previous Acts. It abolished workers' contributions to the compensation fund and introduced differential