DOI: 10.1002/sim.8814

Statistics in Medicine WILEY

ALFRED WATSON MEMORIAL LECTURE

THE STATISTICIAN IN MEDICINE

By PROFESSOR SIR AUSTIN BRADFORD HILL,
C.B.E., F.R.S., Ph.D., D.Sc.
Honorary Member of the Institute

The following is the text of the tenth Alfred Watson Memorial Lecture, which was delivered on 26 February 1962.

In the year 1927 Sir Alfred Watson presented a paper to the Royal Statistical Society. It was entitled *National Health Insurance: A Statistical Review*. Our system of National Health Insurance had then been in operation for nearly fifteen years. Extending to some 16 millions of the population it had been 'accepted generally', in Sir Alfred's own words, 'as an integral part of the social fabric of the United Kingdom' (1). It was, of course, left to a later speaker at that meeting of the Society to point out how much of this acceptance, consequent upon a wise development, was due to Sir Alfred Watson himself, how, indeed, said Dr Heron, it was 'impossible to over-estimate the value of his contributions to this branch of social legislation' (2).

I am, however, at this moment equally conscious of the comment made by another contributor to that same discussion. In asking questions of Sir Alfred or in making suggestions another great actuary, Sir William Elderton, said 'I feel I am rather like a common soldier cross-examining a Field-Marshal' (3). On that scale of values I must myself have been present at that meeting, if not actually in my perambulator, at least at the stage of Cowboys and Indians. But there is something to be said for that; it is an age that has great advantages; it is an *impressionable* age, in other words a good age at which to sit at the feet of a master of a discipline and the leader of a profession, to revere and to learn—to learn a whole attitude and standard of values, not merely a methodology. Having had that experience and that opportunity so long ago, I am indeed grateful for the privilege once again to pay reverence to Sir Alfred Watson in this Memorial Lecture.

Before turning to my subject I would first remind you that of all the strange things that Alice saw in her journey Through the Looking Glass there was one, according to her biographer, that she always remembered most vividly (4).

'...the mild blue eyes and kindly smile of the Knight,—the setting sun gleaming through his hair and shining on his armour in a blaze of light

Hill, A Bradford. Alfred Watson Memorial Lecture. The Statistician in Medicine. *Journal of the Institute of Actuaries.* 1962; **88:** 178–191, reprinted by permission of The Institute and Faculty of Actuaries.

that quite dazzled her—the horse quietly moving about, with the reins hanging loose on his neck, cropping the grass at her feet—and the black shadows of the forest behind.' And then to that quiet and sylvan background, the White Knight's poem, the aged aged man a-sitting on a gate:

'I shook him well from side to side,
Until his face was blue:
"Come, tell me how you live", I cried,
"And what it is you do!"'

I hope I may be held innocent of *lèse-majesté* if I suggest that in instituting and organizing these memorial lectures your Council has been guided by that same scene and that same philosophy.

In one respect I believe they are right. Professionally, and scientifically in particular, we do tend to live in far too restricted an environment. It is both instructive and fascinating to know what goes on elsewhere, even behind the shutters of the house next door—and such proximity might well be regarded as the relationship between actuaries and statisticians. On the other hand, I would gravely doubt whether that instruction and that fascination can be properly provided by those aged aged men who would be better employed sitting on gates. However that may be it is my aim to reveal a little of the life of a statistician in medicine and to discuss in more detail a few of the peculiar problems that life presents.

THE PROBLEMS OF MEDICINE

Medicine can be divided under two broad headings—curative or the care of the sick, preventive or the care of the healthy. Under the former heading we have the fundamental questions that arise in the daily practice of medicine—what is wrong with this patient, what can be done to correct the wrong, what is the upshot likely to be? Diagnosis, treatment, prognosis.

Under the latter heading we have equally fundamental and, really, very similar questions to answer: What is wrong with this population? Upon whom, when and where does this particular ill particularly fall, and why does it so fall? The problems of epidemiology and aetiology, stretching back to the time of Hippocrates. What action can we take to prevent this wrong? The more modern problem of public health.

The answers to all these questions can be sought, it would seem to me, in three different ways. We may devise experiments in the laboratory, using animals as our basic material. We may devise experiments in the everyday world in which we live, using man as our basic material. We may merely observe and record what is happening in man in that everyday world under our eyes. Each of these approaches has its strength and its weakness.

EXPERIMENTS WITH ANIMALS

In an early volume of the *Journal* of the Institute, Professor De Morgan included in his delightful Budget of Paradoxes (5) an account of the derivation of the name of the metal antimony, or, as it came from France, antimoine. The monk in charge of the monastery pigs, he says, observed that they marvellously thrived upon it; and, so, he fed it to his brethren who all died of it. Quite a paradox! In short, we are always faced with the problem of passing from animal to man, and that step may be exceedingly hazardous and difficult. On the other hand although, as Robert Burns said, the best laid schemes of mice and men gang aft agley, we have the advantage in the laboratory that we can repeat an experiment, we can adjust the variables to our liking, we can generally wriggle around to get a clear answer to a specific question. Thereby we may obviously narrow the field of search, of what we need especially to look for in man; alternatively we may strengthen the evidence which is already available in man. But we should remember that lack of such support does not necessarily, or even at all, weaken the evidence available in man himself.

For example, epidemics of influenza have a long, if not respectable, history; but the isolation of the causative viruses was delayed till relatively recently by the fact that the customary laboratory animals have the good fortune to be immune to it. Before that particular door was opened it remained for a laboratory technician to sneeze thoughtlessly at a ferret.

Similarly, the fact that it has not been possible to induce cancer in the respiratory organs of mice by their exposure to tobacco smoke no more contravenes the epidemiological evidence in man than does the production of skin cancers in the same animal by constituents in tobacco smoke prove the dangers of that substance to the respiratory organs of man. We are constantly weighing in the same balance what we know of man, what we know of other species. Often we know all too little of their differences.

In this branch of medical statistical work the major contribution of the statistician lies in the design of experiments and their analysis, experiments, for instance, for the detection of substances in industrial use likely to be carcinogenic, the standardization of drugs and vaccines, the bio-assay of new drugs, the measure of their potency and, perhaps, their toxicity. This last is an essential step—and often over quite a range of animals—before one of the modern wonder-drugs can be let loose on man.

EXPERIMENTS WITH MAN

However, in the last analysis, the value of a drug against a disease can be assessed only by trial against the disease itself, the value of a vaccine can be assessed only by the reaction of the vaccinated when exposed to the disease itself. Thus are we led to my second mode of inquiry—the experiment in man. This, I believe, is our strongest weapon of investigation but one that raises great problems of practicability and, often, of ethics. The strength of the experimental approach lies in our ability to manipulate *at will* the 'guinea pigs' at our disposal, whether they have four legs or two. And it is not so easy to do just that when they have two. The statistician in medicine must therefore not only learn to live with—and, indeed, deliberately to design—experiments that are far from ideal; he must also learn to live a trifle dangerously.

For example, in the *ideal* trial of a new vaccine, we should, I suppose, randomly select members of the general population at risk and randomly allocate them to vaccination or no vaccination and measure the ensuing attack rates from the disease in question. In practice, of course, we have no such powers. We invite *volunteers* to enter the trial and to allow themselves to be randomly allocated to vaccination or no vaccination. In short, we conduct our trial, however meticulously, upon a self-selected group. In such circumstances, as I have written elsewhere (6), 'we may be faultlessly observing the relative incidence of disease in vaccinated and unvaccinated children and both groups may belong, unduly, to small families with careful mothers. Whether the answer we have reached is equally applicable to large families with scatter-brained mothers we do not know, and cannot know.' It is on such a basis, and I believe rightly, that we have developed our public health measures of control of such diseases as pertussis and poliomyelitis. We have learned to pass in defined circumstances from the particular to the general without demanding too exact a correspondence between the two. But in extolling the experimental approach, as is the custom, we should be fully aware of that.

In the same way, in the strictly controlled trial of a new treatment we almost invariably seek to measure its strength and, sometimes, its dangers on a relatively small number of patients in the care of an even smaller number of highly selected doctors. What other approach is possible?

THE MEDICAL STATISTICIAN'S BACKGROUND

Thus, as I said, the statistician in medicine must learn to live a trifle dangerously. But, in this particular field, there is a second lesson that he must learn. Unless himself medically qualified, and thus armed with the weapons of both disciplines, he must learn a great deal of medicine and even, I believe, absorb the doctor's way of thinking. He must, in other words, not only have facility in speaking two languages, he must be able to think in two.

In writing of another profession, these are the words of Lord MacMillan (7): 'So settled are the traditions of honour and fair dealing which the Bar observes and such is the atmosphere which long observance of these traditions has created, that those who have absorbed these traditions and live in that atmosphere acquire unconsciously a sense of what is due to their calling and are scarcely aware of the code of honour which daily guides them, so much has it become their second nature.'

It is my belief that the statistician who is concerned with clinical trials must no less endeavour to acquire that similar code of honour that is second nature to the physician and surgeon. Only thus can he grasp the problems that face the doctor—when can a treatment be withheld, when, sometimes, can a new treatment be given, to the patients in his care?

No medical statistician—and no doctor—can be familiar with the details and intricacies of every disease. But he can be familiar—or make himself familiar as the occasion arises—with the specific diseases involved in trials, just as the medical man can, and should, make himself familiar with the outlines of statistical methodology and the statistician's line of thought. It is a symbiosis that we must seek; if the blind lead the blind shall not they both fall into the ditch?

'It is possible', said the Annual Report of the Royal Statistical Society of 1849, 'to be a very able calculator and yet but a second-rate statist, since, to reach the heights of any science, the first of all qualifications is a just and penetrating spirit.' No doubt. But the corollary of that is an intimate knowledge of just what to penetrate. Without that all is lost.

THE STATISTICIAN IN CLINICAL TRIALS

I give emphasis to this point at this time since some of the clinical workers in the field of clinical trials take a much narrower view of the statistician's role and, indeed, propound a view that I believe to be misleading and dangerous. In this country, at least, they have taken to heart, I am glad to say, Dr Frank Green's aphorism that the statistician should be regarded as an obstetrician and not as a morbid anatomist (8).

The issue lies rather in what they think to be his province when they do come to him first. Thus an anonymous author writes (9) 'it must be realized that the statistician can help only when the questions to be answered are precisely framed by the physician who knows the disease. The type of disease, its course and its treatment, are not subjects to which the former can contribute.'

This theme is developed in detail in Dr Max Hamilton's recently published lectures on the Methodology of Clinical Research (10). Broadly, the question to be answered he defines thus: Does this treatment given in such-and-such a way, by such persons, under such conditions, to patients of a defined type and nature, produce any improvement, as observed and measured by certain defined criteria, and taking into account such factors as age, sex and so on? Accordingly, in planning a trial we have a number of fundamental points initially to settle. Dr Hamilton lists them as:

(*a*) the treatment—what drug, what dose, what system of administration?
(*b*) the patients—what sort, what age, sex and social class?
(*c*) the criteria—what observations, what measurements of the value of the treatment?
(*d*) the control—what factors, such as length of stay in hospital, may make an important difference to the results and thereby need special control?

and finally, with these points settled,

(*e*) the experimental design of the trial.

It is only in the last—in the selection of an experimental design—says Dr Hamilton, that the expert, the statistician, can give advice. 'He cannot tell you', he writes, 'what you mean by treatment and the dose and method of administration; the definition of the population is not his concern. You can test your treatment on senile dements or on adolescent youths; it is for you to decide what you are going to do; he cannot advise. The statistician cannot devise a criterion, for only you, only the clinician, knows what is meant by improvement or recovery. If you do not, nobody else does. So with the identification of the factors to be controlled, the statistician might be able to give you some advice, but not often....

'We need an answer to our investigation as quickly as possible, using the minimum number of patients for the minimum amount of time. This is where the statisticians have the answers. The question of the efficiency of experiments, getting the maximum amount of information from the minimum amount of material, is just what they have been working on, and they know how to squeeze the last drop of juice from your orange. But whether it is a ripe orange or a rotten one, only you can decide.'

If that is to be the role of the statistician in clinical trials, then as far as I am concerned, it is not an orange, ripe or rotten, the answer is a lemon! This is not a symbiosis, it is parasitism and like many forms of parasitism likely to lead to disaster. It may be good advice, indeed, to urge young clinicians to think for themselves; it is not good advice to suggest that they can get effective help in just this one way from an uninformed, and quite possibly misinformed, statistician.

All my experience is to the contrary. There are many ways of thinking and behaving that come naturally to the statistician accustomed to the medical field which by no means come naturally to the medical man. They arise constantly and the medical statistician is aware of them all the time—or should be. To take one or two simple examples that fall under Dr Hamilton's headings, the medical statistician recognizes, and is familiar with the pros and cons of, that difficult question—should a fixed dose be given to all patients in a trial or should it be allowed to vary with the apparent needs of each patient as judged by the clinician?

In many trials the original careful randomization of patients to treatment and control can be later disturbed by selective withdrawals of patients who cease to take a treatment or are proved sensitive to it so that they have to be withdrawn. The experiment is necessarily weakened—indeed we may on occasions have to assess the value of an *intent* to treat rather than a treatment. The medical statistician is aware of these problems and of the effect that they can have upon a trial. He knows too the kind of errors that may be found in the criteria of success or failure of a treatment, whether the criteria are likely to be too insensitive to be of real value at all.

Many of these problems are common to all clinical trials and in designing a specific trial the statistician should know when and how to quiz his clinically expert colleague. Unless he is familiar with them I do not believe he should be so bold, or, perhaps, so naïve, as merely to provide an experimental design to order. This seems to me to be about as futile and dangerous a procedure as providing a formal test of significance to someone else's data of which you know nothing and for which quite possibly a test of significance is the least important test to apply. As Sir George Pickering wrote (11) 'to rely on data, the nature of which one does not understand, is the first step in losing intellectual honesty'.

In all this argument I hope I may be absolved from any desire to make jobs for the statistical boys. Far from it. My one, and very simple, point is that medical statistics is an *applied* science and no one can usefully or safely apply it without an intimate knowledge of the subject-matter. Put at its lowest, in collaborating in clinical trials and other medical problems, we should all be aware of that loyal address to Queen Victoria, the opening words of which were 'Conscious, Your Majesty, as we are of our shortcomings' which led Lord Bowen to propose that most famous amendment 'Conscious, Your Majesty, as we are of each other's shortcomings'.

NATURE IN THE RAW

The third method of study that I listed some time back was the observation of what is happening in man in the everyday world under our eyes. Unfortunately, the medical statistician cannot convince thousands of his countrymen of the contributions that they might make to knowledge (as well as to his own prospect of winning a Nobel prize) if, on an *experimental* basis, they were to eat no animal fats, smoke no cigarettes, take regular exercise and drive their cars more slowly. He has, inevitably, to study what they themselves *choose* to do, in these and other ways, to rely, in short, upon Nature's experiments. The trouble is that Nature is such a poor experimenter. She appears to know rather less of the importance of randomization than today's second-year medical student. Further, she flings the variables around with so prodigal a hand that we are constantly faced with the job of tracing our path to cause and effect through a veritable labyrinth of indirect associations. We have, in other words, to make as sure as may be that extraneous variables do not account for the particular association that we believe to be dominant.

It is certainly more difficult to be sure in the absence of experiment but nevertheless we should keep in mind Cornfield's comment that there is no fundamental difference in the two situations. 'There are merely associations whether observational or experimental that, in a given state of knowledge, can be accounted for in only one way or in several different ways' (12).

The experimentalist's aim is to make two equal groups, groups that do not differ from one another in essential features. With observations we have to ensure that Nature's groups do not differ in essential features. And we shall need to keep in mind that more than one feature of life may well contribute to a specific event. For example, in our present attempts to incriminate dietary factors in the causation of heart disease, we might need to remember what was written a very long time ago— that 'better is a dinner of herbs where love is, than a stalled ox and hatred therewith' (13).

Indeed, just as with many of the infectious diseases—typhoid fever, diphtheria, scarlet fever—there is known to be more than one mode of spread, so with today's chronic diseases it is obviously likely that there will be more than one cause. In most diseases there is ample room for an interaction between heredity and environment. They are not mutually exclusive but both likely, in varying degrees, to make their contribution.

All this it is that makes so difficult (and so entertaining) our disentangling of the chain of causation forged in Nature. Obviously we may often reach inconclusive results. At the best we shall almost invariably be reaching the *most reasonable* explanation of a particular set of facts. We should fully recognize this but at the same time we should fully recognize how exceedingly foolish, and unscientific, is the argument—at present used in particular by those who repudiate the cancer/cigarette smoking evidence—that such evidence is *only statistical.*

STATISTICAL EVIDENCE

Some twenty years ago I was called upon to investigate the mortality from cancer in a group of workmen engaged, in South Wales, in the refining of nickel. The population at risk, workmen and pensioners, numbered about one thousand. During the ten years 1929 to 1938, sixteen of them had died from cancer of the lung, eleven of them had died from cancer of the nasal sinuses. At the age-specific death-rates of England and Wales at that time, one might have anticipated one death from cancer of the lung (to compare with the 16) and a fraction of a death from cancer of the nose (to compare with the 11). In all other bodily sites cancer had appeared on the death certificate eleven times and one would have expected it to do so 10–11 times. There had been 67 deaths from all other causes of mortality and over the 10 years' period 72 would have been expected at the national death-rates. Finally, division of the population at risk in relation to their jobs showed that the excess of cancer of the lung and nose had fallen wholly upon the workers employed in the chemical processes (14).

More recently my colleague Dr Richard Doll has brought this story a stage further (15). In the nine years 1948 to 1956 there had been, he found, 48 deaths from cancer of the lung and 13 deaths from cancer of the nose. He assessed the numbers expected at normal rates of mortality as, respectively, 10 and 0·1.

In 1923, long before any special hazard had been recognized, certain changes in the refinery took place. No case of cancer of the nose has been observed in any man who first entered the works after that year, and in these men there has been no excess of cancer of the lung. In other words, the excess in both sites is uniquely a feature of men who entered the refinery in, roughly, the first 23 years of the present century.

No causal agent of these neoplasms has been identified. Until recently no animal experimentation had given any clue or any support to this, wholly statistical, evidence. Yet I wonder if any of us would hesitate to accept it as proof of a grave industrial hazard? We have the striking contrast between normality in general mortality and in general cancer incidence and the gross abnormality in cancer of just two sites. We have the limitation of that gross abnormality entirely, or almost entirely, to certain workers, those occupied with the chemical processes. We have the apparent disappearance of the abnormality in workers who entered the changed environment after 1923. Yet, let me repeat, all this is statistical evidence of events in a population of whose genetics we are ignorant and whose general environment we have not sought to examine.

Further, we know, of course, that cancer of the lung and cancer of the nose do occur in persons who do not refine nickel. Yet, if you have an interest in maintaining your expectation of life, I would urge you not to tell the refinery workers of South Wales (or more usually their widows) that there can be no case for paying them compensation since the evidence is only statistical.

Nearly twenty years ago Sir Norman Gregg noted, in his clinical practice in Australia, that the mothers of babies who were brought to him with a form of cataract had been attacked by German measles during the early stages of their pregnancy. A world-wide literature has flowed from that astute observation. The effects of the rubella virus upon the eyes, ears and heart of the foetus during its first three months of growth have been widely noted, reported and accepted. The probabilities that some congenital abnormality will follow rubella at specific stages of pregnancy have been measured—at least to a good approximation.

All this is observation. Experiment is, obviously, out of the question. Yet, conceivably, the adult woman who is susceptible to German measles may differ constitutionally or in her environment—for example in social class—from the woman who is immune. The woman who exposes herself to infection in the early months of pregnancy may differ from the woman who sits at home and knits.

In short, there might be other variables that would explain this association between the attack of rubella and the subsequent congenital abnormality. None has been detected or even, I believe, sought; none would seem the slightest bit likely. The purely observational statistical evidence in women points to cause and effect and I know of no one who disputes it.

THE ASSESSMENT OF EVIDENCE

The issue in all such work in the medical field is not that the evidence is merely statistical but whether it is convincing in the sense that one interpretation of the available data is, in the present state of knowledge, very much more likely than any other. *Of course* statistics alone can be a hopelessly insecure and insufficient basis for belief and action. I have not worked in medicine for 40 years without at least learning that. But there are occasions, such as the two I have described, when no scientific worker capable of the simple analysis of data and of logical thought could refuse to accept them as a basis for action. The statistical picture is overwhelming.

I would have thought it unnecessary to lay so much emphasis upon this glimpse of the obvious were it not that the misconception appears so frequently, even amongst scientific workers who should know better.

The question, on the other hand, may well be asked, what does one accept as overwhelming? When does a heap really become a heap? The answer, I submit, is not to be found tidily tucked up in the formulae of tests of significance, useful as they may be. In it there must always be an element of the subjective—the subjective judgment of the particular respondent, of you or me. Indeed, there may be danger in leaning too heavily upon any formal test. For instance, returning to my first example, it is, no doubt, easier to detect a cancer, or other, hazard in industry than in the general population. By the nature of the case the excess incidence is circumscribed in its victims and circumscribed usually in space. However, with many epidemiologically alert minds on the watch in many places the formally 'significant' excess must inevitably turn up by chance from time to time. And it is the 'significant' that will be reported. We shall, therefore, need always to keep in mind the observations of Aunt Ida Clemmens who, you may remember, had premonitions, including the important one that old Mrs Hutchins would not last out the year. Though Aunt Ida missed on old Mrs Hutchins for 22 years she finally made it, and 'if she was right once in 20 times, it proved that she knew what she was talking about' (16). How often did Mr James Thurber hit the nail firmly and squarely on the head, even to recognizing the magic 0·05 level!

Though I fear Aunt Ida would not agree, we shall clearly need to lean heavily upon that scientific safeguard—repetition. How often does the premonition come true? It is axiomatic in science that an experiment should be repeatable, and so described by its author as to be repeatable, by other workers. It is equally true of our statistical observations, our recording of everyday events and their various associations. In the celebrated Brides-in-the-Bath case of my boyhood it was repetition that brought Mr Smith to the gallows. The accidental drowning of an occasional bride can, perhaps, be straightforward clean fun; the accidental drowning of three in identical circumstances is a gross carelessness which does put a strain on the credulity.

The same broad answer in different places, at different times, by different statisticians must obviously be more impressive than the single unsupported observation.

It will be helpful too, as various workers have suggested, if the causation we suspect is biologically plausible. But what is biologically plausible depends upon the biological knowledge of the day. There was no biological knowledge to support (or to refute) Pott's observation in the eighteenth century of the excess of cancer in chimney sweeps. It was lack of biological knowledge in the nineteenth that led a prize essayist writing on the value and the fallacy of statistics to conclude, amongst other 'absurd' associations, that 'it could be no more ridiculous for the stranger who passed the night in the steerage of an emigrant ship to ascribe the typhus, which he there contracted, to the vermin with which bodies of the sick might be infected' (17). And, coming to nearer times, in the twentieth century there was no biological knowledge to support the evidence against rubella.

We are continuously brought back to the fundamental question—what alternative explanation will fit a set of observations, what other differences between our contrasted groups could equally, or better, account for the observed incidences. That is the crux of the matter—and no χ^2 test or other application of the Greek alphabet will answer it. It demands an experience, and acumen, in what to collect, how to seek in the data for essentials, how to interpret. And that, to return 'by the same door as in I went', demands an apprenticeship in the subject-matter. That wisdom comes with the years may be a very dubious proposition but it does imply that one has been around for quite a time. The old man a-sitting on a gate at least had the advantage of a wide variety of knowledge.

Attempts have been made in the U.S.A. to lay down some decisive rules of evidence in this very difficult field (18), rules which will determine our judgment of cause and effect in chronic diseases as Koch's postulates were used to identify the bacterial causation of infective diseases in the last century. I have nothing of that kind to offer you for I am not convinced that we can similarly reduce a wide variety of complex situations to a simple common denominator—that is to one that can be really useful and not merely state the obvious. I am not convinced that we can in advance decide what weight we shall give to this or that piece of evidence.

Take for example (and I quote from Dr Doll (19)) 'the "Kangri cancer" of the skin of the lower abdomen and thighs, which is found among the poor inhabitants of Kashmir who warm themselves in winter by carrying, in front of their abdomen and suspended from the neck, an earthen pot full of live coals (Neve, 1900, 1924) and the "chutta cancer" of the epithelium lining the hard palate which is found in Vizagapatam, India, among persons who are accustomed to holding a home-made cigar with the lighted end inside their mouth, to prevent it from going out (Khanolkar & Surnabai, 1945). These associations have not been demonstrated by detailed retrospective studies of patients and controls nor by prospective studies of persons with and without the specific habits, such as have been reported for a number of other cancers, but a few case histories and the close geographical association between the local habits and the local type of cancer has sufficed to convince nearly all who have been interested in the subject that the relationships are causal.'

As Doll astutely observes, few of us in this country are in the habit of carrying a heated pot close to the abdomen in winter or of smoking our cigars outside in. But would the firm rules of evidence that have been proposed cover such a case? I doubt it.

In general, the statistical-epidemiological inquiry is, by its nature, a 'whodunit', and often an involved 'whodunit'. Surely we must therefore include in the rules every rule that governs the normal everyday weighing and acceptance of evidence? Such inquiry has, I believe, much in common with the law and the law's handling of circumstantial evidence. The production of an orderly and convincing chain of statistical reasoning demands much the same skills. It possesses also much of the same fascination. As Lord MacMillan wrote in his essay on Law and Order (20) 'the mere physical process of arranging, sorting and classifying a miscellaneous heap of objects affords a certain pleasure. An untidy drawer, a cupboard of confused odds and ends, a littered desk rouse in us at least an aspiration—not always, I fear, realized—to put things straight...the pleasure of producing order out of chaos among our possessions, if we really set ourselves to the task, is never-failing.'

To those who would thus seek their pleasures in 'producing order out of chaos' I can confidently recommend the life of the Statistician in Medicine.

REFERENCES

(1) Watson, Sir Alfred. National Health Insurance. A statistical review. *J. R. Statist. Soc.* 1927, **90**, 433.
(2) Heron, David. *J. R. Statist. Soc.* 1927, **90**, 474.
(3) Elderton, Sir William. *J. R. Statist. Soc.* 1927, **90**, 474.
(4) Carroll, Lewis. *Alice Through the Looking Glass*, ch. VIII.
(5) De Morgan, Prof. A. Budget of paradoxes, 1864. *J.I.A.* **11**, 188.
(6) Hill, A. Bradford. The experimental approach in preventive medicine. *J. R. Inst. Publ. Hlth.* 1958, 177–96, 209–18.
(7) MacMillan, Lord. *Law and Other Things*, 1937, p. 196. Cambridge University Press.
(8) Green, F. H. K. The clinical evaluation of remedies. *Lancet*, 1954, **2**, 1085.
(9) Anon. Trials on Trial. *The Glaxo Volume*, 1959, **19**, 18.
(10) Hamilton, Max. *Lectures on the Methodology of Clinical Research.* 1961. Edinburgh and London: E. and S. Livingstone Ltd.
(11) Pickering, Sir George. Disorders of contemporary society and their impact on medicine. *Ann. Int. Med.* 1955, **43**, 919.
(12) Cornfield, J. *Amer. Statist.* 1954, **19**.
(13) Proverbs, ch. 15, v. 17.
(14) Morgan, J. Gwynne. Some observations on the incidence of respiratory cancer in nickel workers. *Brit. J. Industr. Med.* 1958, **15**, 224.
(15) Doll, Richard. Cancer of the lung and nose in nickel workers. *Brit. J. Industr. Med.* 1958, **15**, 217.
(16) Thurber, James. *A Portrait of Aunt Ida.*
(17) Cheever, David W. The value and the fallacy of statistics in the observation of disease. *Boston Med. Surg. J.* 1861, **63**, 449.
(18) Yerushalmy, J. & Palmer, C. E. On the methodology of investigation of etiologic factors in chronic diseases. *J. Chronic Dis.* 1959, **10**, 27.
(19) Doll, Richard (in the Press).
(20) MacMillan, Lord. *Law and Other Things*, 1937, p. 22. Cambridge University Press.