

# The CENTRAL AFRICAN JOURNAL OF MEDICINE

Dr. DAVID LIVINGSTONE

Vol. I. No. 2

MARCH, 1955

## CONTENTS

### ORIGINAL ARTICLES

Pitfalls in Medical Reading	<i>L. J. A. Loewenthal</i> - - -	45
Medical Experiences in Nyasaland	<i>W. H. Watson</i> - - -	54
New Blood Pressure Ranges for Normal People	<i>A. M. Master</i> - - -	62
Age of Childbirth in the Bantu	<i>W. Fraser Ross</i> - - -	69
Silicosis and Carcinoma	<i>R. Paul</i> - - -	71
Clinical Note on Cutaneous Anthrax	<i>P. Mavros</i> - - -	76
Datura Poisoning	<i>J. Thomas and M. Gelfand</i> -	78
Radioactive Isotopes, Part 2	<i>J. F. Duguid</i> - - -	80
African Vital Statistics	<i>J. R. H. Shaul</i> - - -	83

### EDITORIALS

Recognition of Intoxication	86
The Good Doctor	88
New Year Honours	88

### Correspondence—

<i>Obstetrics in the African</i> <i>E. G. Naylor</i> - - -	89
---	----

### African Health—

<i>Opening of Mount St. Mary's Hospital</i>	90
---	----

Obituary — <i>Dr. Ferdinand von Malmborg</i> - - -	90
--	----

Proceedings— <i>Medical Council of S. Rhodesia, Quarterly Meeting</i> -	91
---	----

Book Review - - -	93
Latest Pharmaceutical Preparations	93

# The Central African Journal of Medicine

Volume 1

MARCH, 1955

Number 2

## Pitfalls in Medical Reading\*

BY

L. J. A. LOEWENTHAL,  
M.D., M.R.C.P., D.T.M. & H.,  
Johannesburg.

Mr. President: The honour of being invited from another country to be Guest Speaker at your Annual Congress imposes certain obligations. To mark my appreciation of this great honour I have decided not to choose my topic from my speciality—which, after all, can be of interest to only a few—but to devote my time to a subject which is important to you all, no matter what branch of Medicine you profess. If we regard ours as a learned profession we must be sure that our reasoning is as cogent as that of other scientists, and that such knowledge as we gain is imparted with precision.

A heart specialist recently published his observations on the manner in which patients sat down in his consulting room. To sit down and lean back was termed "sitting at ease"; all other methods, including drawing the chair forward and leaning on the desk, were termed "ill at ease." The author then observed how patients sat when there was no intervening desk, making the comparison under conditions which he thought were adequately controlled. He found a far larger proportion sitting "at ease" and he treated his figures by an approved statistical method to show that this difference was significant. The paper now took on an appearance of great scientific respectability, being seasoned with such expressions as  $X^2=36.8$  and  $P=0.001$ . But what the author forgot, and what you will by now be asking, was a simple question: "What happened to that group of patients who would have leaned forward on the desk when there was no desk to lean on?" As there was no mention of subjects falling flat on their faces, under the new conditions, we must assume that they adopted some other way

of sitting down, possibly including "sitting at ease." Thus at least some of the increase in "sitting at ease" when the desk was removed may have been due to factors which the author neglected. Hence his conclusion, that the presence of a desk between physician and patient may make the latter feel ill at ease, lacks any convincing support. It may be correct, but he has not proved this and the use of statistical methods to analyse his unreliable data is one of the ways in which statistical methods become discredited among the less enlightened.

This elementary example of faulty reasoning is only one of many that can be found by anyone who reads medical literature with a critical eye. It will be missed if, as often happens, only the summary of the article is read or if the article itself is read too hastily. It has led me to consider other common fallacies in medical writing and to describe the ways in which the reader is most commonly confused or misled. Let me state immediately that these harsh words are not directed against the authors' good faith, for most of them seem to be sincere; but many of them are obviously persons of poor general education who show, in particular, the need for instruction in elementary logic. I shall not labour this point, to which I first drew attention nearly twenty years ago (Loewenthal, 1936), but I hope to convince you of its truth as I proceed.

### LANGUAGE

Many authors seem to believe that the use of long words adds to the scientific value of their contributions. Perhaps it is only from a desire to impress, perhaps from more sinister motives: whatever the cause, the result is deplorable, and we often find the simplest thoughts so disguised in verbiage that they become almost unrecognisable. The earliest symptoms of this disease are the tendency to use foreign equivalents of English words; thus "belly" becomes "abdomen," "I have a pain in the neck" becomes "the patient complains of acute cervical discomfort" and instead of being in bed or unable to get up, the patient is

\* An Address delivered at the Annual Congress of the Southern Rhodesia Medical Association, Umtali, September, 1954.

gratified to find that he is "nonambulant." Granted that at the bedside they may be necessary in order to prevent the patient understanding much of what is being discussed, nevertheless the use of these ugly phrases at other times is apt to give the lover of English an acute cervical discomfort. It may be urged that many such expressions provide an economy of phrase; this may be true of some, but certainly not of others: what advantage, for instance, can be claimed for the horrible "prior to hospitalisation" over "before admission?"

I recently saw a newspaper report dealing with the need for simplicity in psychological language (presumably the language used by psychologists was meant). The following was given as an example: "Mary was relatively retarded on digital manipulation in vacuo" and this, according to a clinic worker, meant that "Mary could not waggle her thumb very well."

Let these objections be thought trivial let me remind you that the purpose of language is to convey thought; when the attent on becomes fixed on trying to decipher a message, it is apt to be distracted from the content. Consciously or unconsciously misdirection—the stand-by of the conjuror—creeps in. While our attention is fixed on the right hand, firing a pistol or words of five syllables, the left is producing a bowl of goldfish—or a belief disguised as a fact. When people itch and scratch because they are upset and someone tries to help them by attacking the cause of their upset, the results should make interesting reading; but when these results are produced as "Dynamically-Oriented Brief Psychotherapy: Psychocutaneous Excoriation Syndromes," I, for one, am so exhausted by translating the title into English that I have little energy left for a critical review of the contents. In this particular paper, as a matter of fact, there is no real evidence that the itching and scratching were wholly or partly of emotional origin. We shall meet this type of conjuring trick again when the fallacy of *petitio principii* is considered.

Much more could be said about words and the secondary meanings they have acquired. The use of "significant" can, for instance, be quite misleading. This word has a particular meaning when applied to the statistical treatment of observed facts: its use in other circumstances, for instance in such sentences as "Significant improve-

ment was noted," is of course not illegal, but it may imply more than it states. Words also have their changing fashions: Benjamin (1952) has recently pointed out the unpopularity of "dichotomy" (a perfectly respectable form of classification) and "undynamic" in psychotherapeutic circles. To make good to-day you and your methods must be dynamic, just as a sewing machine must be streamlined, even if it is not going to move at sixty miles an hour against a headwind. Beware, then, of words which imply more than they state; distrust the facile "therefore" and the presumptuous "significant" and regard with suspicion those conclusions which are said to "follow logically" from the observed facts. It is long odds that the author's knowledge of logic is no greater than his readers.

#### LOGICAL LAPSES

Logic has been defined as the science of thought (Creighton, 1910 a); thought is a skilled job and, to give the best results, must be studied and practised as thoroughly as other skilled occupations, from Music to Medicine. Though commonsense is, according to popular belief, all that is needed in everyday life its exclusive use in the sciences is apt to lead the observer astray. Certain rules of thinking have been evolved over the last two and a half thousand years; they have withstood the test of time, and, with few exceptions, still serve to direct the processes of valid thinking. These rules comprise Deductive and Inductive Logic. The former attempts to deduce particular items of truth from universally accepted ones; the latter to increase the sum of human knowledge by deriving universal truths from a collection of specific examples. When you inform the patient's relatives of the hopeless outlook in a case of malignant melanoma with widespread secondaries, you are unconsciously using deductive logic; all cases of malignant melanoma with widespread secondaries are known to die of the disease; this is such a case. Therefore this patient will die of the disease. The possible fallacy that death may occur earlier from some other cause, such as suicide, is irrelevant here. You have deduced a specific truth from a general one.

Inductive logic is more generally used by the research worker. From a specific number of instances he attempts to frame a general law, or a universal truth. Thus an observer who has encountered several cases of pig-

mented cutaneous tumours giving rise to secondary pigmented tumours and invariably causing death may feel confident that such cases carry a poor prognosis. When he reads that from Laënnec's day, all other observers have had the same experience and that the tumour has a characteristic histological structure known as that of malignant melanoma, he may then formulate a law: "that malignant melanoma with widespread secondaries is a fatal disease." You will recollect that it was from this law that the prognosis in a given case was deduced.

Now it is obvious that the great bulk of medical discovery and the vast majority of clinical diagnoses are made by men and women who have not the slightest knowledge of the rules of formal logic. Why then, you will ask, should a knowledge of logic be desirable? I could answer that quite a few discoveries and diagnoses are unsound and would not have been made by persons trained in the habit of logical thinking. This, however, is not my primary object. I propose instead to demonstrate how a knowledge of formal logic may so assist the critical faculty as to help in assessing the value of scientific and pseudo-scientific articles.

One of the principal branches of Logic is the study of fallacies, both in deductive and inductive reasoning. Originally they were no doubt described and classified as an aid to confounding one's opponent in a battle of dialectic, but since then their study has been extended. A fallacy is an error of reasoning (or of interpretation) and familiarity with the commoner fallacies is desirable—I almost said essential—in the critical study of any article which presents a reasoned conclusion. This obviously comprises the greater part of medical literature. Time does not permit me to classify even the commoner fallacies, nor to include examples of the majority. I shall, however, submit a few instances taken from recent articles, purely as illustrations. No reflection is made on the possible excellence of such articles; the facts presented may be of great importance and the conclusions may still be correct, even though they have been derived by faulty reasoning.

The first class of fallacy, which follows naturally from the earlier part of this essay, derives from the use of *Ambiguous Language*. An amusing example of Jevons' is often quoted (Creighton, 1910 b): it occurs in

Chapter XIII of the First Book of Kings, verse 27, where it is said of the prophet: "And he spake to his sons, saying, 'Saddle me the ass.' And they saddled *him*." The italics indicate that the word *him* was supplied by the translators, but a very different meaning may be suggested. A modern example is of far greater importance. There is a class of drugs which act as antagonists of histamine. The liberation of histamine is one of the consequences of some, but not all, forms of allergic reaction; thus the symptoms and signs of urticaria, when due to the liberation of histamine (or a similar substance) may frequently be suppressed by these drugs. Quite incorrectly, these drugs, first called "anti-histaminic," have come to be called "anti-allergic" and are to-day used in enormous quantity for any condition which is thought to be allergic in nature. As far as skin diseases are concerned, it has repeatedly been shown (Warin, 1954) that such drugs are effective only when there is wealing; nevertheless they are constantly being prescribed for other forms of irritating skin disease, such as the eczematous reaction, in which there is no wealing and in which the liberation of histamine plays no part.

A second and similar type of fallacy now comes into view. Some forms of eczema are known to be produced by an allergic mechanism, particularly the contact dermatitis which may follow exposure to plants, cosmetics and chemicals; other eczemas, such as the variety associated with asthma and hay fever, are at least partly dependent on an allergic mechanism. But there is no justification for the belief that *all* eczemas are of allergic origin. Nevertheless it is commonly accepted that "eczema" implies "allergy" and, through the fallacy just explained, "allergy" calls for anti-histaminic drugs.

*Ignoratio Elenchi*. The fallacy of Irrelevant Conclusion is one of the most prolific weeds in the garden of medical literature. Instead of proving what you have set out to prove, you demonstrate some other conclusion which is more or less nearly related. An excellent example is taken from the *Lancet's* report of Question Time in Parliament (1954). It concerns the Minister's failure to re-appoint Dr. Stark Murray to a certain Hospital Board. He was asked the following question: "Can the Minister assure the House that political considerations

did not enter into this, since Dr. Stark Murray is a leading Socialist doctor and his successor is a leading anti-Socialist doctor?" The reply was: "If the hon. Member looks at not one instance, but at all the instances, and at other boards, he will know that there is no foundation for his suggestion." Here is Ignoratio Elenchi at its best. The Minister is asked about Dr. Stark Murray and bases his reply on every other instance, but not on the case of Dr. Stark Murray. Should you still feel that the Minister's reply, if technically illogical, was a reasonable one, then consider the matter in this way: suppose that there was a movement on foot to get rid of socialists occupying important posts. Then obviously the movement must begin by getting rid of one socialist, and the fact that the authorities have not yet got round to dismissing the others is quite irrelevant.

The Irrelevant Conclusion is also seen when personalities are introduced into a discussion (*argumentum ad hominem*); an oft-quoted example is that of the lawyer who was handed a brief marked: "We have no case; abuse plaintiff's attorney." In medical writing this form of the fallacy is rarely seen—possibly because of editorial watchfulness—but it may be thinly disguised, as in the following example. A physiologist described the changes produced in the gastric secretion of dogs subjected to certain unpleasant experiences under controlled conditions. This stimulated an investigator to publish a paper in which the physiologist's experiments and conclusions are severely criticised. I am quite unfit to pass an opinion on the merits of the case and I have quoted it only to point out some fallacies of the investigator: the one we are concerned with here is the phrase, "During the so-called 'experimental periods' . . ." Now why "so-called?" The experiments, as described by the physiologist, may have been good or bad, adequate or inadequate, humane or inhuman; but they certainly were experiments and the use of the adjective "so-called" can have no other purpose than to arouse prejudice against their performer.

The Irrelevant Conclusion is also seen in attempts to transfer the onus of proof, rarely in medical literature, but frequently in arguments and clinical discussions. Here, in effect, the man who cannot prove his point takes refuge in a challenge such as "If you maintain that I am wrong, prove it." This fallacy is often difficult to detect in the heat

of argument, obvious though it may be to the logician. It is so common that Todd (1951) even found it necessary to write an article entitled: "The onus of proving the value of remedies"; the conclusion was that "the advocates, not the opposers, of all remedies should supply proof of their claims."

*Petitio Principii*, or "Begging the Question" is also one of the commoner fallacies. It is first necessary to point out that Begging the Question has nothing to do with Evading the Question, though it is commonly misused in this sense. There are so many ways in which this fallacy can be perpetrated that I propose giving a few concrete examples rather than dilating on the various sub-heads into which it may be classified.

First we may consider the question-begging epithet. If prosecuting counsel refers to the accused as "this bloodthirsty individual" (a procedure which is permitted in certain countries) he obviously assumes the guilt of the accused—the very point that he is trying to prove. In much the same way an author, setting out to prove the emotional or nervous origin of a skin condition, may call it "neuro-dermatitis," or group a number of conditions under the title "the neurodermatoses." I referred earlier in this paper to the use of popular and unpopular words in medicine or, rather, in psychosomatic medicine. You will now appreciate that the adjectives "dynamic" or "undynamic" can become question-begging epithets, expressing approval or disapproval while the subject is still under consideration.

A more complex form of *Petitio Principii* is of frequent occurrence in medical literature: this is known as "reasoning in a circle." A non-medical example given by Creighton (1910 c) will make it clear: "I should not do this act, because it is wrong." "But how do you know that the act is wrong?" "Why, because I know that I should not do it." When we turn to medical literature this variety of begging the question is often found, as for instance in the following example:—

A group of clinicians wanted to find out whether emotional factors played any part in the occurrence of extrasystoles. They took 12 unselected patients and found that "extrasystoles and associated anxiety were observed in these subjects experimentally during a discussion of topics to which they were known to be sensitive or which had previously been associated with extrasystoles." This,

then, is the conclusion. Now let us turn back to Case 3, a hypertensive woman of 51. During a discussion about her mother, during which the patient showed no obvious anxiety, a shower of extrasystoles occurred. I quote the last sentence of this clinical record: "The patient stated that she had not thought the topics discussed disturbing to her *but they evidently touched on matters to which she was sensitive.*" (My own italics.) Why "evidently?" Apparently because they were accompanied by an attack of extrasystoles, and the authors have concluded that extrasystoles are produced by emotional factors. Hence in Case 3 they have assumed the conclusion which they set out to prove.

I trust that you will not be led astray by my frequent use of examples from psychosomatic medicine, and commit the fallacy of supposing that all such studies are fallacious. It is true, indeed, that I have found it easier to demonstrate false reasoning in the writings of psychiatrists than in those of other specialists; but this must not lead us to the next fallacy which I shall discuss, that of False Generalisation. In other words, from the few examples I have submitted, you must not conclude that *all* psychiatric studies employ logical fallacies, nor that *all* psychiatrists dispense with the accepted rules of reasoning. A notable illustration can be found in a recent article by Macalpine (1954) who, herself a psychiatrist, exposes the fallacy of *Petitio Principii* as often perpetrated by other psychiatrists: "Again, some workers start off with the assumption that emotional factors are responsible for a given condition: patients are then investigated to this end, and when such emotional factors or conflicts are found the conclusion is drawn that the original premise was correct, the aetiological conclusions and definitions follow accordingly. A circular argument is thus introduced and continued." Precisely, you will observe, as happened in the episode of the extrasystoles.

This fallacy of False Generalisation is constantly seen in various guises. It is committed whenever conclusions are drawn from an insufficient number of observations, and thus provides one of the many bridges between Logic and Statistics. It is also committed whenever we neglect exceptions and base a universal rule on an overwhelming number of instances. Thus it is to-day customary to regard emotional factors as

the *only* cause of peptic ulcer, though we know that in rare cases other forms of stress—systemic or local—may be incriminated, with equal justification. "The exception proves the rule" say the ignorant, not knowing that "proves" in this context means "tests" (Latin: *probare*), and that where there is an exception there can be no rule. As an amusing example of where faulty generalisation can lead us, let me take you back to this question of peptic ulcer. I have no specialised knowledge with which to argue the question as to whether peptic ulcers are invariably a psychosomatic disease. What I know is that there are people who still hold that these ulcers may arise from other causes, such as faulty diet. Now take this paragraph from the summary of a paper which I mentioned earlier: "The development of this syndrome in young calves prematurely weaned, and the relief of these ulcers by permitting the calves to continue on a milk diet, provides striking evidence for the correctness of the psychoanalytic theory of ulcer formation." Although the author has stated earlier (in a footnote) that the fodder may also be an important factor, he relegates it to a subsidiary place. You see now where the original fallacy has led: by formulating a law through False Generalisation he cannot admit of an exception; hence 98% of weaned calves, between the ages of 12 and 14 weeks, must be showing the effects of mother rejection or some other alteration in cow-calf relationship. The influence of a third party is not mentioned but I feel there is quite a lot of bull around too.

*Inductive Logic*, or the use of observed instances to provide a universal rule, is by its very nature subject to numerous fallacies. The whole question of Cause and Effect, which we cannot now explore, is but one instance of its scope. But for those who wish to reflect on the subject, I recommend the study of three sets of rules, in their diminishing order of exactness: first, we have Koch's Postulates for incriminating a germ suspected of causing a certain disease; next, we have the slightly less rigid, but still exacting, conditions under which a substance may be blamed for provoking an allergic reaction; and lastly we have the criteria by which an illness is to-day considered to be of psychic origin. I regret that time does not allow a closer examination of these and similar problems; a few examples will have to in-

dicating the importance of this field, as well as its vastness.

If one fact obtains, and another invariably follows, it is often expressed concisely as "if p, then q." For example, if a person has syphilis, then his blood shows a positive Wassermann reaction. Now a very common fallacy is to reason thus: "but q obtains, therefore p is true," or, to take our example further, "this patient has a positive Wassermann reaction, therefore he has syphilis." To an audience accustomed to yaws and njoverta, let alone biologically false positive reactions, such a statement sounds ridiculous. Yet I can assure you that it forms the basis of diagnosis in many cases. Even two workers, who recently investigated positive serological reactions in South African natives and doubted whether positive tests were in fact evidence of syphilis, persistently refer to their serological results as the "syphilitic rate."

When we enter the field of Cause and Effect, closer study is a continual source of worry to the intellectually honest. Reflect, for instance, on the history of human malaria: originally residence in swampy regions was considered to be an adequate explanation; later the parts played by the plasmodium and its insect vector were assumed to have solved the problem. Later still the case of the symptomless carrier of the parasite had to be explained, and still more recently the effect of diet on susceptibility to the disease. Thus, in Medicine at least, the demonstration of a single cause or an invariable effect is no longer a matter for serious search. If a man is run over by a bus we do our best to repair his broken bones and crushed tissues; but even in a case as straightforward as this we should, in the consideration of cause and effect, take note of such factors as accident proneness and occupational fatigue. You have all, I am sure, become aware of the growing interest in a connection between cigarette smoking and cancer of the lung. The statistical work, retrospective and prospective, of Doll and Hill (1954) leave one in no doubt that the heavy smoker shows an increased liability to cancer of the lung; but no thinking person, least of all such meticulous workers as Doll and Hill, would be satisfied that the problem of cause and effect had been solved in its entirety. Not only has the question of multiple causation to be considered, but the tempting leap from

statistical evidence to the incrimination of tobacco tar must be viewed with caution. For cigarette smoking connotes not only the inhalation of tobacco smoke, but that of paper, added matter, chemicals that may have been used on the tobacco plant and even, as you well know in this country, the fumes of petrol lighters.

One of the better known fallacies in Inductive Logic is that of False Cause, or "*post hoc ergo propter hoc*." If one event follows another, even with some regularity, we cannot, without further evidence, consider the first to be the cause of the second, for in that case we should have to allow that day is the cause of night. If, then, a young woman tells you that her acne is always worse when she is emotionally upset it is unwise to conclude that the state of her skin is necessarily the result of her psychic state. Both are more probably the result of endocrine changes with water retention, frequently present during the week before the menstrual period. Similarly, I have described the case (Loewenthal, 1954) of a young woman with recurrent inflammation of the eyelids. This happened only when she suffered from a bout of depression and psychiatric treatment was given in vain. Much later it was discovered that she used to fly to the cosmetician whenever she felt an attack of depression was imminent and at these times, and at no others, a coating of nail lacquer was applied. Patch tests showed a positive result and she has had no further attacks of eczema since giving up nail lacquer, though her periodic depression still recurs. Remember that the statement "This patient's rash got better on calcium" means only what it says, that is that his rash got better while he was on calcium; it is usually impossible to establish firmly any causal relationship between the administration of a drug and the improvement noted in a single instance. Only when a series is studied with adequate controls can conclusions be tentatively mooted. I commend to you therefore the scepticism of the Trobriand Islanders, who could not be persuaded that the birth of a child was the result of what had happened nine months before.

#### MISUSE OF FIGURES

Fifteen years ago, when Wilson and I were preparing a paper on statistical methods in Medicine (Loewenthal and Wilson, 1939), we had no difficulty in finding examples of conclusions wrongly drawn from figures

presented. To-day the importance of statistical methods is more widely realised and few papers present their conclusions without some statistical adornment. This usually takes the form of a test which purports to show how often such an agreement or discrepancy would occur through chance alone. If once in twenty trials, the probability ( $P$ ) is 5 per cent., or 0.05; if once in a hundred trials,  $P$  is 1 per cent. or 0.01 and the statistical significance is judged according to this figure. The working out of these statistics calls for no more mathematical knowledge than is imparted to most school-boys, so that many medical research workers perform their own statistical tricks. It often happens, however, that their lack of the necessary training prompts them to employ unsuitable data or inappropriate tests, hence their efforts may remind one of the trained chimpanzee circling the stage on a bicycle—admirable as a spectacle but without getting their performer to any destination.

Other workers invite the help of trained statisticians, with satisfactory results, always provided that the statistician understands the problem and has had a voice in planning the experiment. For when he is simply presented with a series of figures and asked to calculate their significance, his labours cannot possibly cancel out errors committed in the assembling of those figures. He cannot, in fact, make a silk purse out of a sow's ear. The ideal arrangement, of course, is for the research worker to be a competent statistician, as in the case of Doll and Hill, and there is no real reason why medical research workers should not be so trained. But whoever does the statistical work, it must be clearly understood that no refinement of mathematics can compensate for faulty or indefinite data, as I tried to show in the example at the beginning of this paper. This use, or rather misuse, of mathematics was well described by Macaulay, when he referred to the man who uses statistics as a drunkard uses a lamp-post—for support and not for enlightenment. Before we turn to the figures themselves, therefore, and before considering how they may be added, subtracted, multiplied, divided and generally have the square roots battered out of them, let us examine some of the mistakes that can be made in their collection.

White (1953) has recently given a timely reminder of the possible fallacies in the selection of samples for study. We all know

that there is an error inherent in the selection of a sample that is too small and therefore unlikely to be representative; but the converse is by no means a solution, in spite of the popular misconception that the study of the larger and larger series will of necessity bring us nearer to the truth. White has given evidence of two large-scale experiments which turned out to be quite useless as the result of biased sampling, that is the selection of a large series which was not typical of the population as a whole. The use of controls is therefore a subject which calls for more care and knowledge than is usually given to it.

An actual example will make some of these difficulties clearer. Two doctors attempted to assess the connection between certain skin diseases and psychic factors such as Personality Disorder and Emotional Stress. First, a random sample was obtained by selecting every fifth case at Skin Out-Patients for psychiatric investigation. Here then, is the first possible fallacy, that the whole sample, which includes controls, was drawn from a hospital out-patient clinic. We do not know whether Personality Disorder and Emotional Stress in a "random" sample of out-patients is at all comparable with conditions obtaining among the general population; the authors, in fact, postulate that an out-patient clinic *must* contain a larger proportion of psychically abnormal people and hence, with commendable honesty, draw attention to their own sampling error.

Next we must consider a possible source of error which I have not yet mentioned: conscious or unconscious bias in the mind of the observer. In order to obviate this it would have been helpful if the psychiatrist had been in ignorance of the dermatologist's diagnosis, for a patient presenting with the label "neurodermatitis" could not help influencing any psychiatrist who was also a normal, fallible human being. Unfortunately there is no mention of this point and we have no means of telling how much the psychiatrist had gathered about the patient's skin complaint.

The results of this investigation were surprising, for the group of skin diseases which one would normally use as controls—varicose dermatitis, naevi, rodent ulcer, xanthomatosis and so on—showed a high proportion of psychically abnormal subjects. As I have said before, this was explained as being due to sampling errors just short of the point

where it could be said that a man must be crazy to attend a skin out-patient clinic. There are, however, other explanations which depend on the answer to certain questions. What is the degree of accuracy, or what is the margin of error, in results obtained by the psychiatric interview? Does this psychiatrist usually obtain results comparable with those of other psychiatrists? From the studies presented by other investigators, are we justified in accepting as a fact that certain dermatoses are linked with psychic abnormality and, further, that they are the *result* of such abnormality? I believe that many of the discrepancies found so frequently in the results of different investigators could be obviated if more attention were paid to the important and fundamental problem of sampling.

Let us assume that the results of an enquiry have been collected with due regard to the exigencies of sampling, including adequate controls. The next possible source of error lies in the tricks which may be played with these figures.

The illegitimate use of percentages comes readily to mind. It is a useful device and a perfectly proper one, when several unequal series of more than 100 examples in each are to be compared. It is even permissible when some of the series comprise some 70 or 80 examples. But where are we going to draw the line? A recent paper on various methods of treating hypertension provides a good example. Here the series vary from 43 patients to 4 and whereas in the former a response in 20 can, with some justification, be characterised as 47%, in the latter a response in one patient should hardly be regarded as a 25% response. Surely the assumption that, if one patient in four responds, then 25 will respond in every 100 is an example of the fallacy of False Generalisation. Look at it in this way if you are in doubt; a second series of 4 might well show a response in none without causing surprise, and a third series might similarly show a response in two; but the percentage figure would now vary between 0 and 50 and would obviously be useless for purposes of comparison. I have even seen a table in which 1 out of 1 was solemnly entered as 100%. Earlier in this paper I mentioned the conjuror's trick of misdirection; this one must surely rank as the Indian Rope Trick of medical writing.

Let me now use the few minutes that remain to consider the question of statistical devices applied to figures which have been arbitrarily derived. There are many phenomena in medical work which cannot be directly expressed in numbers; thus, whereas a series of red blood cell estimations can be directly entered in a numerical table, the severity of the symptoms of anaemia do not admit of similar treatment. In the one case we have a *quantitative* estimate in the other a *qualitative* one. Now the branch of statistics which I am discussing is concerned with the examination of figures, not the examination of attributes unless these latter are so sharply defined as to give a definite "yes or no" answer. Thus it is permissible to classify a series of cases according to the presence or absence of fits, but further classification into "severe," "moderate" and so on is permissible only if some numerical criterion is introduced; in this way we could separate those patients with ten or more fits a day from those with a smaller number. But there are cases in which no such numerical division can be used. As Yule (1922) says, "The division may also be vague and uncertain: sanity and insanity, sight and blindness, pass into each other by such fine gradations that judgments may differ as to the class in which a given individual should be entered." Thus figures which rely only on the observer's judgment are of a low standard of accuracy, for other observers may well give a different assessment of the qualities under review. It follows that the use of delicate statistical methods to analyse such figures is fundamentally unsound for, as I remarked before, no refinement of mathematics can compensate for faulty or indefinite data.

A single, last example will suffice. In the paper I referred to, entitled "Dynamically-Oriented Brief Psychotherapy: Psychocutaneous Excoriation Syndromes," the author devotes a footnote to explaining how standard deviations and standard errors were derived and how the values of "t" were obtained—all according to normal practice. But when we turn to some of the figures which form the raw material of these calculations, we find the following: "For statistical purposes motivation for therapy was rated according to an arbitrary scale in which poor motivation was scored 1, fair 2, good 3, and excellent 4." I take this to mean that patients who showed no eager-

ness for treatment only scored one point, while those who were as keen as mustard scored four, but I am at a loss to decide how the intermediate gradations were determined. You have only to remember the gross differences observed when different examiners give marks for the same examination papers in order to realise that Seitz's figures are not of the kind upon which statistical analyses should be based. Nor does the employment of impressive words such as "assessment" and "evaluation" add one jot of precision to an indefinite concept.

#### CONCLUSION

This reference to bad language as a disguise for faulty methods brings us back to our starting point. I have drawn attention to some of the pitfalls in the path of the trusting reader of medical literature. Where I have given illustrations of what I consider to be faulty language, logic or calculation, I have tried to do so in the

spirit of Abraham Lincoln: "With malice toward none; with charity for all." And if I have fallen short of this ideal I hope at least to have followed Pope's precept, to "Laugh where we must, be candid where we can,  
But vindicate the ways of God to man."

#### REFERENCES

- Benjamin, J. D. (1952)—*Psychosomatic Med.*, 15, 200.  
 Creighton, J. E. (1910 a)—*An Introductory Logic*. London, MacMillan, p. 1.  
 Id (1910 b)—*Ibid.* p. 169.  
 Id (1910 c)—*Ibid.* p. 181.  
 Doll, R., and Hill, A. B. (1954)—*Brit. Med. J.*, June 26, p. 1451.  
 Elmhirst, E. (1954)—*Brit. Med. J.*, July 10, p. 102.  
 Hall-Smith, S. P., and Norton, A. (1952)—*Brit. Med. J.*, Aug. 23, p. 417.  
 Livesay, W. R., Mayer, J. H., and Miller, S. I. (1954)—*J. Amer. Med. Ass.* 155, 1027.  
 Loewenthal, L. J. A. (1936)—*East Afr. Med. J.*, Oct., p. 1.  
 Id (1954)—*The Eczemas*. Edinburgh, Livingstone. p. 198.  
 Loewenthal, L. J. A., and Wilson, W. A. (1939)—*Brit. Med. J.*, July 15, p. 110.  
 Macalpine, I. (1954)—*Modern Trends in Dermatology*. (Ed. MacKenna, R. M. B.), London, Butterworth, p. 22.  
 Question Time (1954)—*Lancet*, May 15, p. 1035.  
 Todd, J. W. (1951)—*Lancet*, Sept. 8, p. 438.  
 Warin, R. P. (1954)—*Brit. Med. J.*, May 8, p. 1066.  
 Yule, G. U. (1922)—*An Introduction to the Theory of Statistics*. London, Charles Griffin, 6th Ed., p. 8.



This work is licensed under a  
Creative Commons  
Attribution – NonCommercial - NoDerivs 3.0 License.

To view a copy of the license please see:  
<http://creativecommons.org/licenses/by-nc-nd/3.0/>

This is a download from the BLDS Digital Library on OpenDocs  
<http://opendocs.ids.ac.uk/opendocs/>