

## CIGARETTES, CANCER, AND STATISTICS

*Sir Ronald Fisher*

SEVEN OR EIGHT years ago, those of us interested in such things in England heard of a rather remarkable piece of research carried out by Dr. Bradford Hill and his colleagues of the London School of Hygiene. We heard, indeed, that it was thought that he had made a remarkable discovery to the effect that smoking was an important cause of lung cancer. Dr. Bradford Hill was a well-known Fellow of the Royal Statistical Society, a member of Council, and a past president—a man of great modesty and transparent honesty. Most of us thought at that time, on hearing the nature of the evidence, which I hope to make clear a little later, that a good *prima facie* case had been made for further investigation. But time has passed, and although further investigation, in a sense, has taken place, it has consisted very largely of the repetition of observations of the same kind as those which Hill and his colleagues called attention to several years ago. I read a recent article to the effect that nineteen different investigations in different parts of the world had all concurred in confirming Dr. Hill's findings. I think they *had* concurred, but I think they were mere repetitions of evidence of the same kind, and it is necessary to try to examine whether that kind is sufficient for any scientific conclusion.

The need for such scrutiny was brought home to me very forcibly about a year ago in an annotation published by the British Medical Association's Journal, leading up to the almost shrill conclusion that it was necessary that every device

of modern publicity should be employed to bring home to the world at large this terrible danger. When I read that, I wasn't sure that I liked "all the devices of modern publicity," and it seemed to me that a moral distinction ought to be drawn at this point. There is the attitude of a man (may I say, I think it is an entirely rational attitude and one within his own competence to judge) who says, "There seems to be some danger—I can't assess whether it is infinitesimal or serious. This habit of mine of smoking isn't very important to me. I will give up smoking as a kind of insurance against a danger which I am quite unable to assess." That seems to me a perfectly rational attitude. What is not quite so much the work of a good citizen is to plant fear in the minds of perhaps a hundred million smokers throughout the world—to plant it with the aid of all the means of modern publicity backed by public money, without knowing for certain that they have anything to be afraid of in the particular habit against which the propaganda is to be directed. After all, a large number of the smokers of the world are not very clever, perhaps not very strong-minded. The habit is an insidious one, difficult to break, and consequently in many, many cases there would be implanted what a psychologist might recognize as a grave conflict.

If there is cause for fear, let there be warning. But there is no reason for this in the first rational response that I described—that does not require scientific proof that there *is* reason to fear. There is only the possibility that there is reason.

Before one interferes with the peace of mind and habits of others, it seems to me that the scientific evidence—the exact weight of the evidence free from emotion—should be rather carefully examined. I may say, I am not alone in this. I have been interested to note that leading statisticians in this country also—and I contact a good many statisticians both in my own country and here—are exceedingly skeptical of the claim

that decisive evidence has been obtained. In the popular press, the matter seems to be argued, as always, a little off the simplest lines. For example, I find people saying, "These statisticians think this"—"These statisticians think that," or representing that this kind of evidence which has been produced has been attacked as being merely statistical. Now I should be the last person to attack evidence for being merely statistical, because for a great part of my work I have been concerned with the problem of *how* experimentation should be carried out, *how* reasoning processes should be applied to the data supplied by experimentation or by survey so as to give really conclusive answers.

Progress has been made during the last twenty-five years. A large part of the educated world, at least in the statistical field, has become aware that, by taking certain specific precautions, entirely unchallengeable conclusions can be obtained in the experimental field. The work was done primarily in agriculture, where problems of experimentation attracted the attention of leading agronomists at an early time. The key words which emerged in the course of these inquiries—replication, randomization, and control—are now widely understood.

We understand that replication is required for two purposes: it is necessary in order to add precision to our results by diminishing the error to which they are subject, and it is essential in a more important way, as supplying the only means of the estimation of such error.

Although replication is essential in this way, it is not sufficient without the added precaution of randomization, that is, the assignment of the different treatments—which may be manurial treatments, or different varieties of agricultural crops, or different methods of tillage—to the plots set aside for the purpose, in such way at random as to guarantee the validity of the experiment, and in particular of the estimate of error to which it is subject. This necessity for randomization was

brought home to agriculturists largely because it was found that human judgment was very liable to err in this matter, that if one tries to think of numbers at random, one thinks of numbers very far from at random. If one tries to think of a card of an ordinary playing deck, it's well known (perhaps it's not so well known—it is known to me, at least) that red cards are thought of more readily than black cards, that odd numbers are thought of more readily than even numbers, and that the Queen of Diamonds is a hot favorite. This proclivity of the human mind affects any consciously guided choice or assignment of material. Agriculturists, at least, do not trust themselves to choose plots and say that they have been chosen at random. They use decks of cards or, more expeditiously, in recent years, some of these large collections of random sampling numbers which some of you may have seen at the ends of books of tables and perhaps wondered what on earth they can be for. They are in constant use in the design of experiments.

There is a logical aspect, too, of randomization which needs emphasis in this connection. Supposing we have an association—an observable and verifiable association—between two things. I remember Professor Udny Yule in England pointing to one which illustrates my purpose sufficiently well. He said that in the years in which a large number of apples were imported into Great Britain, there were also a large number of divorces. The correlation was large, statistically significant at a high level of significance, unmistakable. But no one, fortunately, drew the conclusion that the apples caused the divorces or that the divorces caused the apples to be imported. The early logicians would say that *post hoc* is not the same as *propter hoc*, or in other words—as it would be put in the early years of our century, when statisticians had had perhaps ten years' experience of the correlation coefficient as a means of research—that *correlation is not causation*. The fact is that if two factors, *A* and *B*, are associated—clearly, positively,

with statistical significance, as I say—it may be that *A* is an important cause of *B*, it may be that *B* is an important cause of *A*, it may be that something else, let us say *X*, is an important cause of both. If, now, *A*, the supposed cause, has been randomized—has been randomly assigned to the material from which the reaction is seen—then one may exclude at a blow the possibility that *B* causes *A*, or that *X* causes *A*. We know perfectly well what causes *A*—the fall of the dice or the chances of the random sampling numbers, and nothing else.

But in the case where randomization has not been possible, these other possibilities lie wide open and should be excluded, or at least every effort should be made to exclude them, before we can assert that causation has been established. When I spoke to Bradford Hill in the early days of this affair, he was entirely unwilling to claim that causation had been proved. He said he didn't see what else it could be, but he was certainly unwilling to make the claim which is being made vociferously during the last year or two by committees reporting to the Medical Research Council in England, and to the American Cancer Society. Now, randomization is totally impossible, so far as I can judge, in an inquiry of this kind. It is not the fault of the medical investigators. It is not the fault of Hill or Doll or Hammond that they cannot produce evidence in which a thousand children of teen age have been laid under a ban that they shall never smoke, and a thousand more chosen at random from the same age group have been under compulsion to smoke at least thirty cigarettes a day. If that type of experiment could be done, there would be no difficulty.

The principles of experimentation—which, as I mentioned, were developed in the agricultural field, where the need for them was greater or more manifest—have spread, and spread rapidly and healthily, into the other experimental sciences. And I suppose during the last fifteen years a dozen important books have been written on the design of experiments, prin-

cially to make clear what these principles are in their particular applications in chemistry, physics, biology, or what you may will.

But the most difficult field for the application of these principles has always been the medical field. This is partly because you can do things to a rat or rabbit which may not be good for it, feeling that in a good cause you have a right to do so. But no one feels—and especially a medical man could not feel—that it is right to do things to a human being which probably will do him harm. Consequently, deliberate experimentation has not been very widely used in the medical field. There is a movement at the present time to organize clinical trials, let us say, of new drugs or of new antibiotics in such a way that an impartial judgment of comparing the new with the old may be obtained by hospital staffs. And that would involve applying the new and the old at random to some of the hospital patients. So long as no body of medical opinion can say with confidence that one is better than the other, or perhaps that in matters usually as complicated as this, for what cases one drug is the better and for what cases the other—so long as that state of ignorance remains, it would be perfectly fair, I think, to clear the air by such simple experimentation.

But manifestly we cannot experiment with the same freedom that is possible with agricultural animals and laboratory animals in other sciences. For lack of that, medical research has had to rely a good deal on uncontrolled experiments, uncontrolled observations; and of course from the time of Jenner onwards there were numerous cases where an observant (and also, I may say, an experimental) physician may be able to make out an exceedingly strong case. Jenner's work was not completely passive. And Dr. Snow, who studied and in the end quelled the occurrence of cholera in London, used a very large number of different types of inquiry in order to gain sufficient confirmation of his important con-

clusion, namely, that it was fecal contamination in the water supply that was responsible for the cholera, an opinion that is easy to take for granted at the present time, but which in the absence of any knowledge of the organisms concerned—or, indeed, knowledge that the disease *was* caused by an organism—was a considerable advance, just as Jenner's was also in the case of smallpox. Consequently, when inconclusive evidence is criticised on the grounds that it is inconclusive, it is not uncommon for medical men to defend it, perhaps with certain indignation, on the ground that in the past medical science has made notable advances primarily—not solely, never only, but primarily—by the observational method.

Now, in the sciences we also have cases in which experimentation is impossible. In astronomy, for example, experimentation, you might say, has only just begun. And in those sciences we must use what I may call *sidelights*.

Let me illustrate this possibility with a very few instances. The first reports of Hill and Doll made a very simple claim. They said that the additional amount of lung cancer observed in patients was proportional to the amount of tobacco they consumed. That simple conclusion was quite rapidly withdrawn, and it was admitted that tobacco consumed in the pipe or in the cigar did not appear to have so close an association with lung cancer as that consumed in the cigarette. And this was a puzzling thing. After all, tobacco is burned in all three cases. The effluvia, smoke, or aerosol from the burning tobacco passes into the mouth, partly into the throat, partly, indeed, into the lungs, in all three cases. It is not obvious—it is not what one would guess at first sight, it was not what Doll and Hill guessed at first—that the one sort of smoke should be comparatively or perhaps wholly innocuous and the other sort should have the effect of inducing the beginnings of a dreadful disease.

And now I must go back and recall just what the kind of evidence it was that Hill and Doll laid before us at the be-

ginning, and in what ways it has been extended by other evidence.

The first inquiry was to take about 1500 patients in a number of different hospitals who had been diagnosed as suffering from lung cancer. Of course the diagnosis is enormously aided in recent times by the use of radiology. The lung cancers can be perceived by their shadows when X-rays are passed through the lungs. Consequently there was good reason to think that these patients—although they were alive and had not been examined post-mortem—really were lung cancer cases. Arrangements were made to record their smoking habits and their smoking history: non-smokers, cigarette smokers, pipe smokers, estimates of the amount of daily consumption of tobacco in each case, and a number of other questions. A similar number, perhaps a few more, of non-cancer patients from the same hospitals received the same questionnaire, and the comparison between these two samples, one of them selected as being lung cancer cases and the other as being in hospitals from some other condition, was made of the classification by smoking habit. And it appeared from that that the cigarette smokers were more common among the sufferers from lung cancer than they were among other patients, and that within the cigarette smokers, heavy cigarette smokers were more common among the lung cancer patients than medium or light cigarette smokers.

The statement that consumers of tobacco in other forms were associated with lung cancer seems to have largely evaporated. I should say a word about it because it represents a common cause of error in statistical investigations, namely, the kind of error which flows from the difficulty of a perfect classification. Everyone can make a rough classification of cigarette smokers or pipe smokers or non-smokers, but there will be borderline cases. There are people who, though they may prefer a pipe when they have the opportunity, yet may be constrained by duress, such as in the intervals of a play when



there is very little time, to smoke a cigarette. There are also distinguished and expensive restaurants, as well as aircraft, who don't like the customer to pull out a pipe. Consequently there is an overlap in the practices and habits of different people; there may not be exactly the same interpretation put on the questionnaire by all the different subjects; and, in fact, a good many pipe smokers may be classified as cigarette smokers, and vice versa. There is bound to be some mixture of the classes in any inquiry on a complicated question. And so the first results did seem to show some effect on pipe smokers and cigar smokers, but it is quite clear that the amount was much smaller than was at first thought, and certainly no more than might easily arise due to misclassification. At least it would be very foolish of anyone who wished to make a case for saying that cigarette smoking was a cause of lung cancer to bring in the evidence about pipe and cigar smoking.

When an unexpected discrepancy occurs, it is a common reaction (I won't say, a failing—it's part really of the scientific discussion which data deserves) to think up some reason for it. This, in effect, may be something like what the logicians would call a "special pleading." That is to say, the making of an assumption, which might be true, which might, indeed, not be true, but which, if true, would help to explain what is otherwise inexplicable. For example, the cigarette contains paper, or, rather, is contained by paper. One doesn't smoke paper much in pipes. There are, indeed, special papers supplied to pipe smokers who wish to enjoy their tobacco in that way. But most pipe smokers and, I suppose, all cigar smokers, do without paper. And it could be, therefore, that it's the consumption of paper that is the really dangerous practice. Then, also, it has been observed that the temperature at which the tobacco is burned is higher in the case of the cigarette than in the case of the pipe, and, it could be (though it certainly is not known to be) that burning at a higher temperature is a

condition for producing something quite unknown, something quite unexplored, something quite hypothetical, in the tobacco smoke which would be capable of producing lung cancer. It is also known that the tobacco used as pipe tobacco and for cigars is more thoroughly fermented before use than is that used in cigarettes, or at least in the predominant source of cigarette tobacco, in Virginia. I think those who prepare the tobacco produced in Virginia are rather acutely aware, that the price per pound is high, there is loss of weight in fermentation, and it is as well not to lose 10 per cent more weight than is necessary. And so, on the whole, the Virginia tobacco is rather lightly fermented. You could imagine—you could claim even—as a special pleading, that it was the unfermented condition of the Virginia tobacco, largely used in cigarettes, that was responsible for the supposedly noxious fumes which the burning of such tobacco produces. Discussion is full of such things.

One of the first people in the United States that spoke to me on the matter, a lady, said, "Of course, cigarette smokers inhale; pipe smokers don't." And of course she laid her finger on an extremely important point. Cigarette smokers in this country, I believe, generally inhale. In England, some do and some don't. When I was a little boy, it was thought that smoking was all right and did you no harm, but inhaling was perhaps a perverse practice and might not do you any good. And so, at any rate my generation, and perhaps some decades of younger men, had a certain amount of warning against this particular practice. I imagine it is something like that that explains the difference in practice between the two countries.

Now, Doll and Hill, in their first inquiry—the one that I've gone over approximately—*did* include in their questionnaire, which was put both to the cancer patients and to the patients from other diseases, the question: "Do you inhale?" And the result came out that there were fewer inhalers among

the cancer patients than among the non-cancer patients. That, I think, is an exceedingly important finding. I don't think Hill and Doll thought it an important finding. They said that probably the patients didn't understand what inhaling meant. And what makes it far more exasperating, when they put into effect an exceedingly important research, based on the habits of the medical profession, by asking about 60,000 doctors in Great Britain to register their smoking habits, and about 40,000 of them did so cooperatively, I am sorry to say that the question about inhaling was not in that questionnaire. I suppose the subject of inhaling had become distasteful to the research workers, and they just wanted to hear as little about it as possible. But it is serious because the doctors could have known whether they were inhalers or not; they could have known what the word meant; perhaps they would have consulted each other sufficiently to lay down a definition which the rest of us could understand. At any rate, there would have been no alibi if the question had been put to a body of 40,000 physicians.

So, our evidence about inhaling is embarrassing and difficult. There is no doubt that inhaling is more common among heavy cigarette smokers than among light cigarette smokers in Great Britain, where inhaling is not nearly a universal practice. There is no doubt that cancer is commoner among the heavy cigarette smokers than among the light cigarette smokers. Consequently, if inhaling had no effect whatever, you would expect to find more inhalers among the cancer patients than among the non-cancer patients. There would be an indirect correlation through the association of both with the quantity smoked. Now, of course, in what was reported everything was thrown together; and yet, in the aggregate data, it appeared that the cancer patients had the fewer inhalers than the non-cancer patients. It would look as though, if one could make the inquiry by comparing people who smoke the same number of cigarettes, there would be a

negative association between cancer and inhaling. It seems to me the world ought to know the answer to that question.

Before I stop, in fact, I hope I shall make clear that there is a case for further research, and I shall only mention two areas which would seem to be profitable for investigation. I would stress the importance of what could be done comparatively easily with rather little expense, namely, to ascertain unmistakably what the facts are about inhaling. If inhaling is found to be strongly associated with lung cancer, it would be consonant with the view that the products of combustion, wafted over the surface of the bronchus, might induce a pre-cancerous and thence a cancerous condition. But if there is either no association at all or a negative association, we should have to reject altogether that simple theory of the causation of cancer. The subject is complicated, and I mentioned at an early stage that the logical distinction was between *A* causing *B*, *B* causing *A*, something else causing both. Is it possible, then, that lung cancer—that is to say, the pre-cancerous condition which must exist and is known to exist for years in those who are going to show overt lung cancer—is one of the causes of smoking cigarettes? I don't think it can be excluded. I don't think we know enough to say that it is such a cause. But the pre-cancerous condition is one involving a certain amount of slight chronic inflammation. The causes of smoking cigarettes may be studied among your friends, to some extent, and I think you will agree that a slight cause of irritation—a slight disappointment, an unexpected delay, some sort of a mild rebuff, a frustration—are commonly accompanied by pulling out a cigarette and getting a little compensation for life's minor ills in that way. And so, anyone suffering from a chronic inflammation in part of the body (something that does not give rise to conscious pain) is not unlikely to be associated with smoking more frequently, or smoking rather than not smoking. It is the kind of comfort that might be a real solace to anyone in the fifteen years of

approaching lung cancer. And to take the poor chap's cigarettes away from him would be rather like taking away his white stick from a blind man. It would make an already unhappy person a little more unhappy than he need be.

For my part, I think it is more likely that a common cause supplies the explanation. Again, we do not know. I do not put forth any explanation as proved, but as requiring investigation. The obvious common cause to think of is the genotype. We are all different genotypes. I suppose in this nation there must be well over 150 million different genotypes. If one studies cancer in mice (and I suppose about half the mice of the world are kept to study cancer with), if one examines any of the many (and there are thousands) of inbred lines of mice (where we can get a hundred or two hundred individuals of the same genotype to study)—if you take, then, any two such lines of differing genotypes, they will, I believe, invariably be found to differ in the frequency, in the age incidence, and in the type of cancer which those mice suffer from. Consequently if there is any genotypic difference between the different smoking classes, we may expect differences in the type or frequency of cancer that they display.

That is the second line of research which I should like to advocate, a little bit more difficult than that which is concerned with inhaling, but certainly well within the capacity of modern methods in human genetics. It certainly could be ascertained, as a matter of fact, whether in the different smoking classes of nonsmokers, cigarette smokers, pipe smokers, cigar smokers (the minor classes, perhaps, of snuffers and chewers perhaps might not be sufficiently numerous, but in those first main four classes it could certainly be ascertained) whether there was evidence that they differed genetically. It wouldn't be a long shot to guess that they did. After all, we choose these things for ourselves. I know that there are families in which there would be some pressure on a growing boy or girl to be a nonsmoker because his father and mother

firmly believe that smoking is an objectionable habit, or perhaps an irreligious habit. But most of us choose for ourselves, and even though one may have been exposed to opportunities—temptations, if you like—to smoke cigarettes from a fairly early boyhood, it is not uncommon to find people who never smoke anything but a pipe. Why? Because they are made that way. They are the sort of men who take to the pipe and don't take to cigarettes, just as there are other men who would never take to a pipe but constantly feel the need of cigarettes. It is not, then, a very long shot to guess that there is a genetic component which distinguishes the different smoking classes. And that is the second piece of research which I think is extremely urgent.

I have criticised the over-confidence shown at least in public utterances or published reports of anonymous committees on this subject, and I do not suppose that Bradford Hill, at least, is at all to blame for that overconfidence. The worst effect of that overconfidence, so far, is that it seems to have held back the various teams of workers. They are well supplied with money—the Medical Research Council is not stinting money on cancer research, and the American Cancer Society is obviously exceedingly well supplied with money. And yet, I think nothing but overconfidence that they had found the solution, that they had the game in the bag, could have prevented them from following up some of the other lines of inquiry which are much needed. I have said nothing, for example, so far of the very striking fact that at the same level of cigarette smoking, dwellers in towns have considerably more lung cancer than dwellers in the country. I don't know any extensive piece of research which has been set on foot to get to the bottom of that important difference.

The desire to make a strong sensation, to bring home the terrible danger to these passive millions, has led writers to stress the very alarming fact that lung cancer is a disease increasing, one of the few important diseases that are increasing

in frequency. It is not so important in the United States as it is in England, but it is an important cause of death in both countries. It has been increasing over the last fifty years. It is frightening. But it shouldn't be used to frighten people.

The change over recent decades gives not the least evidence of being due to increasing consumption of tobacco. We can't tell much about the absolute magnitude of this secular change. It is certain that radiology has facilitated the detection of lung cancer enormously, that radiological apparatus and radiologists are much more abundantly available for our populations than they formerly were. I do not know that there are not remote and secluded communities where patients with lung cancer are not looked at by radiologists, but that proportion of our populations must be still decreasing. Again, the attention of the medical profession has been forcibly drawn to lung cancer, and it invariably happens that when the attention of the medical profession is drawn to any disease, that disease begins to take up more space in the official reports—it is more often seen and more often diagnosed with confidence; death certificates more often include that particular disease. Consequently it is not easy to say how much of the increase is real. I think part of it must be real, because there's no doubt that the populations concerned have been enduring or enjoying a very considerable increase in urbanization. The big metropolitan cities have been growing rapidly. In England, smaller towns have been running together into extensive masses called conurbations, like those of Clydeside or Merseyside or the Birmingham region. Even in the country, even in what used to be remote villages, there are motorbuses regularly which take the young men and women into cinemas perhaps six or eight miles away. You might say that the whole population during the last twenty, thirty, forty years has been becoming steadily urbanized, and as the urban rate for lung cancer is considerably greater than the rural rate, in my country as in yours, we must recognize here the possibility

of one real cause of the increase in lung cancer. There may be others.

But the only good comparison we can make in respect of the time-change is that between men and women. The same apparatus, the same radiologists, the same physicians diagnose both men and women. Whatever effects improved apparatus may have, whatever effects an increased attention to the disease may have, will be the same in the two sexes. Whatever effects urbanization may have you would think might be the same in the two sexes. Consequently, we can, at least, inquire whether the rate of increase of lung cancer in men is the same, or greater, or less, than the rate of increase of lung cancer in women. For it is certainly true, I think in both our countries, that whereas the smoking habits of men have not changed very dramatically over the last fifty years, yet the smoking habits of women have changed a very great deal. And on making that comparison, it appears that lung cancer is increasing considerably more rapidly—absolutely and relatively—in men than it is in women, whereas the habit of smoking has certainly increased much more extensively in women than in men. There is, in fact, no reasonable ground at all to associate the secular increase in lung cancer as has been done with dramatic eloquence, I suppose as part of the campaign of bringing home the terrible danger, just as though it was impossible that statistical methods of inquiry should supply a means of checking that very rash assumption.

And so I should like to see those two things done, one immediately and quickly: an inquiry into the effects of inhaling, and secondly, a more difficult but certainly a possible task of seeing to what extent different smoking classes were genotypically conditioned. And I believe that only overconfidence, if it is allowed to have its way, could prevent those further inquiries from being made.