International Journal of Epidemiology

Vol. 4, No. 3
Printed in Great Britain

Popper's Philosophy for Epidemiologists

CAROL BUCK[1]

Buck, C. (Dept. Epidemiology and Preventive Medicine, University of Western Ontario, London 72, Canada). Popper's philosophy for epidemiologists. *Int. J. Epid.* 1975, **4**: 159–168.
This paper discusses the application of Popper's philosophy to epidemiological research, examining in particular the problems of replication without risk of refutation, of mistaking statistical sophistication for deductive inference, and of dealing with causality at a general level. An example is given of a Popperian approach to the test of a causal hypothesis concerning cancer of the cervix.

INTRODUCTION

Bryan Magee (1) opens his book on Popper with the remark that *Karl Popper is not, as yet anyway, a household name among the educated.* Nor is his a household name among scientists, few of whom are interested in the philosophy of scientific inquiry. Although many have been given an introduction to the logical foundations of science during their training, only a minority renew the acquaintance. This is unfortunate, particularly at present when there are signs of an impending withdrawal of the *carte blanche* that for so long has been offered to science. We hear protests against the lavish financial support given to investigators among whom only a few are capable of creative work. The suggestion has been made more than once that mediocre scientists no longer be given even modest research funds and that instead they be put to work in a supporting capacity with demonstrably creative investigators. I think it possible that among the apparently mediocre there is a significant number who could be creative were they aware of Popper's philosophy of science. Productive habits of thought might be acquired if thought were given more emphasis and action less.

Platt (2) says that one of the outstanding differences between productive and unproductive branches of science is the use of strong inference in the former. By this he means the orderly progression from hypothesis to prediction and then to observation as a way of discarding and revising hypotheses. Although Platt mentions Popper only in passing, it is Popper's philosophy that he prescribes. *How many of us write down our alternatives and crucial experiments every day, focusing on the* exclusion *of a hypothesis? . . . We become 'method-oriented' rather than 'problem-oriented'. We say we prefer to 'feel our way' toward generalizations. We fail to teach our students to sharpen up their inductive inferences. And we do not realize the added power that the regular and explicit use of alternative hypotheses and sharp exclusions could give us at every step of our research. The difference between the average scientist's informal methods and the methods of the strong inference users is somewhat like the difference between a gasoline engine that fires occasionally and one that fires in steady sequence. If our motor engines were as erratic as our deliberate intellectual efforts, most of us would not get home for supper* (2).

Slater (3) has summarized Popper's philosophy for the benefit of investigators in psychiatry, a field in which the line between art and science is particularly difficult to draw. The difference between the failings of psychiatry on the one hand and those of the biological medical sciences on the other is interesting. Bad science in psychiatry usually arises from the creation of a global hypothesis which is poetically elaborated upon by its creator and his followers without any appeal to facts that would be capable of refuting it. Bad science in medical biology usually arises from an investigator behaving as a naturalist and applying technically ingenious methods to the collection of observations that test no hypothesis. For both such tendencies Popper supplies a powerful remedy.

Since my purpose is to discuss the application of Popper's philosophy to epidemiology, I shall make no further gratuitous comments about the failing of other branches of medical science. What shortcomings do we have that attention to Popper might correct?

[1] Professor of Epidemiology and Preventive Medicine, University of Western Ontario, London 72, Canada.

Epidemiologists are exceptionally concerned about their method of approach. In few other medical sciences is so much attention devoted to the philosophical, as opposed to the purely technical, aspect of method. The reason for this is that in epidemiology the experiment plays a relatively minor role. It is usually necessary to create a quasi-experiment out of naturally occurring phenomena. The dangers inherent in a quasi-experiment have made epidemiologists especially attentive to the logical foundations of their work. But their concern with logical rigour is narrow in a Popperian sense and could be raised to a more creative level if Popper's views were better known. Had I encountered Popper's writings earlier, I would have done many things differently.

A SUMMARY OF POPPER'S PHILOSOPHY*

The traditional view of science has been that induction, the formation of a hypothesis based upon observation, is the mainstay of the scientific process. This has always presented a logical difficulty because a hypothesis so derived is forever vulnerable to denial by the first observation that proves an exception. If by induction we believe that all swans are white, our belief can be overthrown by one appearance of a black swan. Popper has rejected the traditional view and regards induction as a dispensable concept. He believes that all observations are made to test some hypothesis already in mind, the derivation of the hypothesis being a work of the imagination. One uses deductive reasoning to make predictions from the hypothesis and thus to state what it prohibits. Scientific discovery in Popper's view is based solely upon a hypothetico-deductive process and advances by disproof rather than proof.

Sherlock Holmes said that *when you have eliminated the impossible whatever remains, however improbable, must be the truth.* This aphorism nearly summarizes Popper's philosophy, except that 'must' in the last phrase should be altered to 'may'.

Given that knowledge is advanced only by testing hypotheses and discarding those that fail, the value of a hypothesis depends upon the degree to which testable and thus refutable predictions can be made from it. A useful hypothesis, therefore, has two important characteristics:

(i) The phenomena that it predicts are open to observation. In fact, if a hypothesis lacks this characteristic it lies outside the boundary of science.
(ii) It is rich in empirical content and so identifies many phenomena that would be incompatible with it.

If a new hypothesis is advanced in competition with one currently held, we should prefer the new hypothesis if at least one of the following criteria is satisfied:

(i) It makes more precise predictions;
(ii) It explains more of the previous observations;
(iii) It explains the previous observations in more detail;
(iv) It has passed tests which the older hypothesis has failed;
(v) It has suggested new tests (i.e. made new predictions) not suggested by the older hypothesis;
(vi) It has unified or connected phenomena not previously considered to be related.

The first, fourth and fifth criteria are most specifically Popperian. The second, third and sixth are familiar as components of the general tenet of scientific parsimony. All are superior to the vague advice often given to scientists that the 'simplest' theory should be preferred.

These criteria emphasize the need to advance beyond the purely *ad hoc* hypothesis that explains a phenomenon equally well but no better than an older hypothesis. For as Popper says, it is always possible to produce a theory to fit any set of observations. We advance only when we suggest refutable consequences of the theory. As long as this requirement is met, there is nothing wrong in forming an *a posteriori* hypothesis because a prior hypothesis was completely contradicted by the observations made to test it. In an epidemiological study of the influence of season of conception upon the IQ of the offspring, Knobloch and Pasamanick (4) observed a relationship almost exactly the opposite of the one that their theory predicted. They were criticized (5) for immediately switching hypotheses in order to explain what had been observed. The criticism is to some extent justified, but only because they put forward few suggestions for refuting their new hypothesis.

* The reader who wishes to go beyond this summary can turn to any one of three books by Karl Popper: 1. *The Logic of Scientific Discovery*, 2nd ed., Harper & Rowe, New York, 1968; 2. *Conjectures and Refutations*, 4th ed., Routledge & Kegan Paul, London, 1972; 3. *Objective Knowledge: An Evolutionary Approach*, Oxford University Press, London, 1972.

To those who find Popper's approach congenial, it may seem that such a view of science is self-evident and that Popper can hardly be credited with having created a new philosophy of science. Medawar in *The Art of the Soluble* (6) raised this question and pointed out that William Whewell of Trinity College, Cambridge, had expressed as early as 1840 a view very similar to Popper's when he stated that the failure to refute a hypothesis *after trying to do so* was much more significant than the mere absence of known facts that were incompatible with it. Similarly, Lord Acton (7) in 1895 stressed the importance of considering objections to a theory: *Remember Darwin taking note of only those passages that raised difficulties in his way; the French philosopher complaining that his work stood still because he found no more contradicting facts; Baer who thinks error treated thoroughly, nearly as remunerative as truth, by the discovery of new objections; for as Sir Robert Ball warns us, it is by considering objections that we often learn.* Coming closer to our own field, we have the following statement from Karl Pearson (8): *The assumption which lies at the bottom of most popular fallacious inference might pass without reference for it is obviously absurd, were it not, alas! so widely current. The assumption is simply this: that the strongest argument in favour of the truth of a statement is the absence or impossibility of a demonstration of its falsehood.* Thus Popper has distinguished forebears. It is he, however, who has described the hypothetico-deductive method in fullest detail and has given the most rigorous logical reasons for adopting it.

Just at a time when Popper's philosophy has been attracting more general attention, there appeared in the popular scientific literature the work of Edward de Bono (9). His essays on creative thinking provide a useful companion to the work of Popper. To search for all the refutable consequences of a hypothesis demands highly imaginative thinking. Imagination is needed to arrive at the hypothesis in the first place, let alone to suggest rigorous tests for it. Although imaginative thinking is commonly regarded as an inborn ability which cannot be taught, de Bono suggests many techniques for enhancing whatever abilities a person has received from nature. His Popperian approach is well illustrated by the following quotation from *The Use of Lateral Thinking*:

> *Every decision is made with some degree of uncertainty. Confidence in a decision does not depend on the lack of any alternative, for that might only indicate lack of imagination, but on the ability to see many alternatives all of which can be rejected. In the making of any decision it can be useful to use one's own or someone else's lateral thinking to generate alternative views so that rejection of these can strengthen the decision. A free-thinking devil's advocate, instead of casting doubts on a good decision, can only succeed in strengthening it.*

De Bono also offers this appealing example of the untestable *ad hoc* hypothesis: *Sects which assemble on mountain tops on predicted days of doom to await the end of the world do not come down on the morrow shaken in their ideas, but with a renewed faith in the mercifulness of the Almighty.*

Popper's tenets are more compelling to the intellect than to the emotions. The creator of a hypothesis is inclined to have a personal attachment to his creature and is not intuitively disposed to seek eagerly for a means of refuting it. Platt has suggested that it is easier to follow Popper's criterion of falsifiability if one makes multiple hypotheses to explain a phenomenon. Indeed, as Lakatos emphasizes (10), the importance of Popper's philosophy lies in its use of refutability to choose which among competing hypotheses should be pursued rather than to reject hastily any single theory. De Bono makes the further point that the exercise of developing multiple hypotheses improves one's capacity for imaginative thought and thus increases the likelihood of arriving at a highly innovative hypothesis.

The importance of breaking the habit of clinging to and endlessly buttressing a single hypothesis is richly illustrated by research on cigarette smoking. Here hypotheses are embraced with such fervour that the thought of refutability becomes heresy. Yerushalmy (11) deserved great credit for his ingenious attempt to test the causal significance of smoking in low birthweight by comparing babies born before and after their mother took up the smoking habit. Yet his readers have paid far less attention to the ingenuity of his approach than to the confounding effect of maternal age which made his results inconclusive. I think that their attitude can be explained by the exasperation created among epidemiologists by ignorant and sometimes corrupt defenders of smoking. In such an emotional climate, it is very easy to abandon Popper's cardinal precept that refutation is heuristically more powerful than confirmation.

THE APPLICATION OF POPPER'S PHILOSOPHY TO EPIDEMIOLOGY

In Popper's terms, the formation of some hypotheses, however preliminary, precedes the collection of data on the distribution of a disease. The data must be capable of refuting the hypothesis. By judicious oscillation between the making of hypotheses and the gathering of data, epidemiologists should be able to discard many hypotheses without recourse to experiment. The experiment, always a most difficult undertaking in epidemiology, is reserved for testing the hypothesis that has survived ingenious deductive efforts to falsify it in observational studies.

Modern epidemiology, concerned with all kinds of illness and not just with communicable disease, is a relatively new branch of medical science. Consequently its aims are not always well understood. How frequently we hear the request to *give us the epidemiology of disease X*, by which is meant an account of the time trend of X and the way in which its frequency is related to age, sex, geographical area, occupation and other demographic variables. I fear that too often this is all that is expected of epidemiology. A commonly held view is that the epidemiologist gathers data from which other scientists will build hypotheses. This view excludes epidemiologists from the exciting process of scientific deduction, and puts them into the business of induction which Popper so convincingly dismisses as irrelevant.

Some epidemiologists appear content with this arrangement. Some get around it by working collaboratively with other scientists so that they may participate in the deductive process. Some, especially those who are medically qualified, go directly to the literature of another branch of medicine in order to construct their own hypotheses. A recent example of the last can be found in the work of Anderson (12). He reached his hypothesis about the role of dietary antioxidants in myocardial infarction by studying the literature on muscular dystrophy, realizing that the latter might be analogous to the myocardial response to ischaemia.

The epidemiologists who are content to serve as preliminary data-gatherers should not be criticized, as long as they are there by choice and not by a mistakenly diffident view of the role of epidemiology in medical science. The two avenues of escape open to an epidemiologist who is not content with data-gathering are probably equally acceptable. What matters is that epidemiologists be fully aware of the part that they could play in the scientific process.

Three features of epidemiological research are particularly worthy of examination from Popper's point of view.

1. *The problem of replication without risk of refutation*

The collection of data is an indispensable part of the epidemiological process, but it can become a wasteful activity if it is not inspired by a deductive aim.

If an epidemiological study in one country has shown that a disease is related to a particular set of variables, why should another epidemiologist repeat the study in his own country? There are only two legitimate reasons. The first has to do with the need to confirm observations because the play of chance can create spurious associations. The role of chance can never be eliminated, but can be given a measure of probability from tests of statistical significance with whose logic epidemiologists are better acquainted than most medical scientists. A large probability that chance explains an association may derive from a paucity of observations. Here, then, is one justification for replication of an epidemiological study. When this is the purpose, the replication should be as exact a reproduction as is possible of the original study. Wilson (13) has criticized epidemiologists for failing either to undertake or to publish an exact reproduction when it is called for.

The second reason for replication is deductively much more powerful in Popper's sense. Suppose one hypothesizes that an endogenous degenerative process plays a major causative role in a disease. A positive association of the disease with age would support the hypothesis. But the identification of a population in which age had no relationship to the disease would suggest the alternative hypothesis that it is caused by prolonged exposure to an exogenous factor, present in some populations and absent in others. The greater one's age, the longer would be one's exposure to the factor if it were present. Repetition of a study that showed age-dependency thus should be made in a population as different as possible from the original one. Replication then becomes an attempt at refutation. If the study is repeated in a population little different from that in which the association with age was first observed, no serious attempt is being made to refute a hypothesis. After many confirmations of the positive relationship between hypertension and age, it was rather belatedly realized (14) that in some populations the relationship is absent. With an earlier search for such exceptions and a

vigorous pursuit of the alternative hypothesis that they suggest, our knowledge of the causes of hypertension might now be further ahead.

Thus an epidemiologist must consider carefully his reason for repeating the work of others. Is the reason a statistical one, to decrease the role of chance in an association, or is it a deductive one, to identify circumstances under which the association does not exist? Sometimes the latter has already been accomplished but has gone unheeded. Consider for example, the well-known positive relationship between latitude and multiple sclerosis, which has been variously interpreted as reflecting the importance of sunshine, diet or race. Japan's low rate was regarded as a tolerable inconsistency until Poskanzer and his colleagues (15) seized upon this as a basis for a new hypothesis, namely that poliomyelitis in the adult is a precursor of multiple sclerosis. The new hypothesis would explain Japan's aberrantly low frequency of multiple sclerosis in terms of its previously 'tropical' pattern of poliomyelitis (widespread infection in childhood and subsequent immunity in adults).

The epidemiological investigation of coronary artery disease offers another example. Many studies have been made, in very similar populations, of the physiological and psychological correlates of heart disease. The number of nearly identical investigations now exceeds the statistical requirement for replication of their common findings. Many of them have been undertaken without any intent to test predictions based upon competing hypotheses. Yet their discordant findings with respect to the association between coronary artery disease and socio-economic status could have led much earlier to the formation of new hypotheses, such as the interesting one that it is status incongruity rather than socio-economic level that plays a causal role in coronary disease (16).

The wastefulness of epidemiological research that has no deductive impetus has been emphasized in a recent paper by Neri, Hewitt and Schreiber (17), in which they discuss the relation between water hardness and mortality. The title itself is interesting, *Can Epidemiology Elucidate the Water Story?* They refer to four classes of hypotheses concerning the apparently protective effect of hard water. The first category deals with a purely indirect role of hard water arising from its association with other factors related to mortality. The second deals with bulk protective ingredients in hard water, the third with trace protective elements and the fourth with harmful factors in soft water. They say that it is inconceivable that one research design would provide a basis for discriminating among all four categories and that epidemiologists should forsake their hopes of bringing off a scientific *coup* and be content to ensure that their projected studies have at least some discriminatory power among the four kinds of hypotheses. After reviewing fifty studies of water hardness in relation to mortality, they conclude that half had no discriminatory power at all and that the rest were mainly concerned with discriminating between a general and a cardio-specific effect of hard water. They finish with the remark that *epidemiological studies will have something to contribute if conducted . . . with due regard for specifying and testing explicit hypotheses.*

Although Popper's philosophy requires that a new hypothesis should explain previous observations, it demands more than this. If an epidemiologist tests his hypothesis against existing observations and finds it compatible with all of them, he cannot stop there without falling into the trap of *ad hoc* reasoning. Yet the exercise of reviewing a hypothesis in the light of previous observations is useful because it provides an inexpensive preliminary exposure to refutation. The proper use of this exercise is illustrated by Matthews and Glasse (18). After testing against all the facts their hypothesis that the tropical disease Kuru is infectious rather than genetic in origin, they made a prediction about the future incidence of Kuru in New Guinea after cannabilism has ceased. On the fulfilment of this prediction their hypothesis now rests.

Epidemiologists are sometimes asked why John Snow invariably receives the greatest prominence in any historical account of their subject. The simple reason is that he was successful in establishing a causal mechanism for an important disease and that he accomplished this by methods that we now call epidemiological. What needs more emphasis in our praise of Snow is the consistently deductive nature of his approach. When he gathered new data, it was always for the purpose of testing a highly specific prediction of his hypothesis. The incident of the Broad Street pump in Soho has become famous as one of the first experiments in epidemiology (19). Snow's recommendation that the pump be closed may not have been viewed by him as an experiment so much as a mandatory action. Having shown that people living beyond the area of the pump who had used its water were attacked, while residents of the area who did not drink from it were spared, what other action could have been recommended? In

my view, Snow should be revered more for his analytical epidemiology than for his famous experiment.

2. *The problem of mistaking statistical sophistication for strong inference*

Because epidemiologists are sophisticated in analytical methods, it is very easy for them to believe that they are doing better than most medical scientists in following Popper's principle of falsifiability. I refer here to the epidemiological craft of dealing with confounding and intervening variables, which has been well expounded in Susser's book, *Causal Thinking in the Health Sciences* (20). However, one must realize that there is a vital difference between the usual exercise of this craft and Popper's principle of testing a hypothesis by searching for its refutable predictions.

The separation of confounding variables is an important technique for determining whether we have a new hypothesis or whether we are merely restating a previously established relationship. It can also be useful for analysing observations that have been made to test the predicted consequences of a new hypothesis. But since it is only a technique it should be the servant not the master of our imagination. This point needs to be put forcefully to students of epidemiology who can easily be led by their teachers to believe that the analytical methods of epidemiology are an end in themselves.

An excessive recourse to statistical devices for tidying up epidemiological data may cause us to miss the exception that in Popper's sense tests the rule. Miettinen (21) has recently shown how by standardizing and summarizing stratified data we can obscure important modifiers of cause and effect relationships.

Adherence to Popper's principle of starting with a hypothesis will protect the epidemiologist from the pitfalls of multivariate analysis. In the absence of a specific hypothesis it is all too easy to emerge from such an analysis with a set of near-zero regression coefficients. This gives the false impression that something has been refuted when in fact all that has been accomplished is the statistical formulation of a tautology. If, however, a multivariate procedure is undertaken to test a carefully formulated hypothesis, we are more likely to put into the equation variables that are conceptually distinct from one another and therefore amenable to statistical separation.

The epidemiologist's preoccupation with methodological niceties may sometimes cause the premature rejection of a hypothesis that should be pursued. Fliess (22), with a viewpoint different from mine, mentions Raymond Pearl's suggestion in 1929 that the protein of the tubercle bacillus might be curative of cancer. Pearl's suggestion arose from his observation that cancer and tuberculosis were rarely associated. Because his observation was based on autopsy data it was dismissed, perhaps too hastily, as an example of Berkson's fallacy. Yet now, nearly 50 years later, BCG is being tested as a treatment for cancer. The irony of this cannot escape even the most methodologically compulsive epidemiologist.

3. *The problem of dealing with causality at the general level*

In the epidemiological study of chronic disease, work usually begins at such a broad level of causality that the predictions of a hypothesis have to be couched in vague terms. Consequently the observations capable of refuting the predictions may be difficult to specify. Unless we realize this, we are in danger of applying the principle of falsification at the naive level against which Lakatos (10) warns the recent convert to Popper. If the predictions of a theory fail to be upheld by observations that are themselves inappropriate, no useful refutation has been accomplished. Let us again take multiple sclerosis as an example, this time with the measles virus as a supposed causal factor. Few of us would think it appropriate to regard the absence of measles antibodies in some patients as a phenomenon prohibited by the hypothesis and therefore as a test of its validity. More likely an epidemiologist would postulate that other factors were also capable of inducing multiple sclerosis. A more appropriate test would then be based on the prediction that epidemics of measles would be followed after some defined interval by an increase in the frequency of multiple sclerosis. But this prediction could lead to difficulty if several common viruses were causally involved. A rise in multiple sclerosis due to measles might be obscured by concomitant decreases in the incidence of the other viral infections. With some knowledge of immunology, the epidemiologist might hypothesize that the antigenic properties of a set of viruses, to which measles belongs, provoke an aberrant immune response that may result in multiple sclerosis. The prediction could then be restated in terms that would permit a more meaningful refutation.

In their textbook of epidemiology, MacMahon and Pugh (23) say that a causal hypothesis based upon epidemiological evidence is strengthened if

it invokes some known, or at least possible, physiological mechanism. Although their book contains many Popperian ideas, they would be even more in tune with Popper if they said that a physiological mechanism is useful in suggesting a means of refuting the hypothesis.

Another example of the difficult early stages of epidemiological research can be drawn from psychiatry. Much work has gone into the search for environmental causes of mental illness, with the putative causes expressed in only the most general terms. If a prediction is made at this level in order to test a hypothesis, its refutation may have little value. Take for example the remarkable variability in the results of inquiries into the relationship between loss of a parent in childhood and subsequent mental illness. Such variability could easily lead to spurious refutation of a hypothesis that depended upon finding the relationship. Birtchnell (24) explains the variability by noting that there are many possible consequences of bereavement and that these undoubtedly vary according to other features of the bereaved child's situation. This and other examples of global variables lead him to a conclusion that is essentially a hopeless one for epidemiology, but not a surprising one in view of the title of his paper: *Is There a Scientifically Acceptable Alternative to the Epidemiological Study of Familial Factors in Mental Illness?* I do not share his view that epidemiologists must be confined to the study of global variables. If they appreciate the need, surely they can refine a global variable into its deductively more powerful components.

Cassel (25) offers an excellent illustration of the need to make a hypothesis as rich as possible, even if its components are rather global in nature. He cites a study in which it was predicted that life-stress would increase the risk of complications among pregnant women who lacked the social support necessary to deal with stress. The study confirmed the prediction. But had either of two simpler hypotheses been formed their predictions would have been refuted, because neither life-stress nor social support was related to complications of pregnancy when considered alone.

The foregoing examples take me back to a point touched upon earlier. Can the epidemiologist, unless he specializes in a particular disease, construct hypotheses for which productive refutation is possible? It is easier to specialize than repeatedly to find expert collaborators as one moves from disease to disease. The latter may be preferable, however, from the standpoint of creativity. Returning to de Bono (9), we hear from him convincing arguments for not specializing. He believes that a creative hypothesis is more likely to occur to a person not so saturated by conventional ways of regarding a phenomenon that he is unable to look at it upside down and backwards.

There is nevertheless an advantage to be gained from the very general level of causality at which epidemiologists frequently operate. At this level it may be possible to apply another of Popper's criteria, namely that a strong theory unifies or connects hitherto unrelated problems. Cassel (25) took this viewpoint when he suggested that epidemiologists might be more productive if they worked sometimes from cause to disease rather than in the more usual reverse direction. He suggested, for example, that the rapid urbanization and industrialization of a rural population might produce stress that enhanced susceptibility to many diseases. I think this approach has been discouraged by the distaste of epidemiologists for a putative causal factor that 'causes too much'. In an attempt to generalize from Koch's postulates for establishing causality in microbial disease, it has been said that one should be wary of a factor that appears to play a part in many diseases. If this advice had been taken seriously at the time when the various symptoms of lead poisoning were regarded as separate diseases, an appreciation of the toxic nature of lead would have been greatly delayed. As Susser (20) suggests, we would be more faithful to Koch's real meaning if we predicted the specific cause-effect relationships embodied in a 'super-cause'.

AN EXERCISE IN THE POPPERIAN APPROACH TO EPIDEMIOLOGY

Let us look at the history of the epidemiological investigation of cancer of the uterine cervix. In 1948, Kennaway (26) published his painstaking compilation and imaginative interpretation of data concerning cancer of the uterus. He concluded with the following statement: *The data collected in this paper suggest the existence of two factors which may increase the incidence of cancer of the uterus, namely:*

(1) *A factor which is opposed by the Jewish practice of abstention from intercourse during most of the first half of the ovulatory cycle, and*
(2) *A factor intensified in both married and single women by descent in the economic scale.*

As the investigation of cancer of the cervix continued, it became clear that the age of beginning

sexual intercourse was another factor, with the risk of cervical cancer elevated among women who began intercourse under the age of 17.

In 1968, 20 years after Kennaway's paper, Rawls (27) advanced evidence in favour of the hypothesis that a venerally transmitted virus (Herpes type 2) is the carcinogenic agent.

The infectious hypothesis is compatible with the social class gradient of cancer of the cervix. It is also compatible with the fact that the disease is more strongly associated with the number of sexual partners than with the frequency of sexual intercourse, a point made by Terris (28) when he discussed infectious versus chemical carcinogenesis as a cause of cervical cancer.

The infectious hypothesis has been made to embrace the association with age at first intercourse by appending to it one of the following:

(a) the adolescent cervix is more susceptible to viral infection;
(b) the adolescent cervix is more susceptible to viral carcinogenesis;
(c) first coitus under age 17 is associated with promiscuity and therefore with high exposure to infection.

The last fits poorly with Terris' finding that women with multiple sexual partners were at greater risk of cervical cancer only if they married after the age of 17.

By what explanation can the infectious hypothesis be reconciled with the remarkable sparing of Jews? The protection against infection afforded by circumcision is an unlikely explanation because the importance of circumcision in cancer of the cervix is negligible among non-Jewish women. Martin (29) has suggested that the endogamy practised by Jews creates a barrier to the entrance of an infectious agent into Jewish populations. Or, an explanation based upon racial immunity to infection might be invoked. The last explanation would be difficult to accept in the face of the equally low frequency of cervical cancer among Ashkenazi and Sephardic Jews in Israel (30). These two groups of Jews have been geographically separated for a long period of time, they differ in certain anthropometric indices, and they have a markedly different incidence of a genetic disorder, Tay-Sachs disease (31).

As an exercise in Popperian epidemiology, I shall restate the infectious hypothesis so that it unites the observations about age at first intercourse and abstention from intercourse by orthodox Jews during much of the pre-ovulatory phase of the menstrual cycle. The revised hypothesis is this: susceptibility of the cervical epithelium to viral carcinogenesis is greatest in the first few years after menarche and in the pre-ovulatory phase of the menstrual cycle, because at both times the cervix lacks the protective effect of progesterone. At the menarche, anovulatory cycles are common and are only gradually replaced in the ensuing 2–3 years by ovulatory cycles in which progesterone counteracts oestrogen-induced changes in the cervical epithelium.

The next step in the exercise is to make as many testable predictions as possible from the re-stated hypothesis. I have made the following eight predictions:

1. Cancer of the cervix among Jewish women should occur mainly among those who do not observe the orthodox sexual ritual.

Stewart *et al.* found that among Jewish women, fewer cases than controls abstained from intercourse during the first week after the menses (30). The proportion stating that they observed the full sexual ritual (Laws of Niddah) was not, however, greater in the controls. In Israel, Pridan and Lilienfeld found no difference between cases and controls in the woman's own assessment of her orthodoxy in religion (32).

2. The beginning of sexual relations at an early age should be a more important risk factor in cervical cancer among orthodox Jewish women than among all other women, because the sexual ritual of the former affords protection once ovulatory cycles have been established.

This prediction receives some support from Stewart *et al.*, whose data show a much greater difference between cases and controls in median age at first intercourse among Israeli Jewish women than among New York Jewish or non-Jewish women (30). On the other hand, Pridan and Lilienfeld found only a small difference between their Israeli cases and controls in the proportion beginning intercourse under the age of 20 (32).

3. Among orthodox Jewish women who did not begin sexual activity at an early age, there should be no social class gradient in the incidence of cervical cancer. After ovulation has been established, the protection afforded by the Jewish sexual ritual would override any differences in hygiene associated with social class.

Pridan and Lilienfeld (32) found little evidence of a social class trend in their case-control study of

Jewish women, although they did not examine the relationship according to age at first intercourse.

4. Among orthodox Jewish women, the number of marital partners should not be an important risk factor in cervical cancer, because the Jewish sexual ritual would make exposure to infection irrelevant.

In Pridan and Lilienfeld's study, the number of marital partners was not important, except among Jewish women emigrating to Israel after 1955 (32). If the more recent immigrants were less orthodox, this observation would be consistent with the prediction. But if exposure to infection can be increased by the sexual activity of either marital partner, the prediction is not supported by their observation that husbands of cases had more sexual partners than husbands of controls.

Nor is this prediction fully supported by data from the New York-Israel study showing number of marital partners according to age at first marriage (30). Among women marrying at age 20 or later, the difference between cases and controls in the number of marital partners was smaller among Jewish women in Israel than among Jewish women in New York. However, the difference was equally small among non-Jewish women in New York.

5. In a case-control study of women who began coitus at an early age, the age at menarche should be later among the cases in order for a higher proportion of their sexual activity to have occurred during the phase of anovulatory cycles.

Stewart *et al.* (30) provide some data to test this prediction. The proportion of women beginning intercourse under the age of 20 is given according to age at menarche. The prediction is not borne out, because the difference between cases and controls in age at first intercourse was not at a maximum among women with late menarche.

6. The beginning of sexual relations before the age of 17 should be less strongly related to cervical cancer among women who bore children prior to that age, because their average age of establishing ovulatory cycles will have been earlier. I have found no data for testing this prediction.

7. A history of repeated spontaneous abortion should be associated with cervical cancer, because susceptibility to abortion is to some extent related to low levels of progesterone. The validity of this prediction rests on the assumption that the progesterone level in pregnancy is correlated with that during the normal menstrual cycle.

Rotkin (33) provides some evidence against the prediction. He found no difference between cases and controls in the overall frequency of abortion. But if habitual aborters are only a small proportion of all women with at least one spontaneous abortion, his observation does not entirely refute this prediction.

8. The incidence of cervical cancer in various Jewish populations should be highly correlated with their degree of orthodoxy.

I have put this prediction last because unless one makes bizarre assumptions about the extent of Jewish orthodoxy in the United States, it is resoundingly refuted by the equal incidence of cervical cancer among Jewish women in Israel and New York (30). Had this prediction been considered first, there would have been little point in continuing the exercise.

Although the exercise has supplied no clue to the epidemiological puzzle of cancer of the cervix, I hope that it has illustrated Popper's approach to the construction and demolition of a hypothesis.

ACKNOWLEDGEMENTS

I would like to thank Dr. Kathleen Stavraky and Professor Carl Robinow for their helpful comments on the first draft of this paper.

REFERENCES

(1) Magee, B.: *Popper*. London: Fontana/Collins, 1973.
(2) Platt, J. R.: Strong inference. *Science* **146**: 347, 1964.
(3) Slater, E.: The psychiatrist in search of a science. *Brit. J. Psychiatry* **122**: 625, 1973.
(4) Knobloch, H. and Pasamanick, B.: Seasonal variation in the births of the mentally deficient. *Amer. J. Pub. Health* **48**: 1201, 1958.
(5) Sterling, T. D.: Seasonal variations in the birth of the mentally deficient? *Amer. J. Pub. Health* **50**: 955, 1960.
(6) Medawar, P. B.: *The Art of the Soluble*. London: Methuen, 1967.
(7) Acton, Lord: *A Lecture on the Study of History*. London: Macmillan, 1911.
(8) Pearson, K.: *The Grammar of Science, Part I*. London: Adam and Charles Black, 1911.
(9) de Bono, E.: *The Use of Lateral Thinking*. Harmondsworth: Pelican Books, 1974.
(10) Lakatos, I.: Criticism and the methodology of scientific research programmes. *Proc. Aristotelian Soc.* **69**: 149. 1968.
(11) Yerushalmy, J.: Infants with low birth weight born before their mothers started to smoke cigarettes. *Amer. J. Ob. & Gyn.* **112**: 277, 1972.

(12) Anderson, T. W.: Nutritional muscular dystrophy and human myocardial infarction. *Lancet*, **2**: 298, 1973.

(13) Wilson, E. B.: A critical look at statistical epidemiology. *Cancer* **16**: 510, 1963.

(14) Truswell, A. S., Kennelly, B. M., Hansen, J. D. L. and Lee, R. B.: Blood pressures of Kung bushmen in Northern Botswana. *Amer. Heart J.* **84**: 5, 1972.

(15) Poskanzer, D. C., Shapira, K. and Miller, H.: Multiple sclerosis and poliomyelitis. *Lancet* **2**: 917, 1963.

(16) Shekelle, R. B., Ostfeld, A. and Paul, O.: Social status and incidence of coronary heart disease. *J. Chron. Dis.* **22**: 381, 1969.

(17) Neri, L. C., Hewitt, D. and Schreiber, G. B.: Can epidemiology elucidate the water story? *Amer. J. Epid.* **99**: 75, 1974.

(18) Matthews, J. D. and Glasse, R.: Kuru and cannibalism. *Lancet* **2**: 449, 1968.

(19) Snow, J.: *Snow on Cholera—A Reprint of Two Papers.* New York: The Commonwealth Fund, 1936.

(20) Susser: M. *Causal Thinking in the Health Sciences.* New York: Oxford University Press, 1973.

(21) Miettinen, O.: Confounding an effect modification. *Amer. J. Epid.* **100**: 350, 1974.

(22) Fliess, J. L.: *Statistical Methods for Rates and Proportions.* New York: John Wiley, 1973.

(23) MacMahon, B. and Pugh, T. F.: *Epidemiology, Principles and Methods.* Boston: Little, Brown & Co., 1970.

(24) Birtchnell, J.: Is there a scientifically acceptable alternative to the epidemiological study of familial factors in mental illness? *Soc. Sci. & Med.* **8**: 335, 1974.

(25) Cassel, J.: Psychosocial processes and 'stress': theoretical formulation. *Int. J. Health Services* **4**: 471, 1974.

(26) Kennaway, E. L.: The racial and social incidence of cancer of the uterus. *Brit. J. Cancer* **2**: 177, 1948.

(27) Rawls, W. E., Tompkins, W. A. F., Figueroa, M. E. and Melnick, J. L.: Herpesvirus Type 2: Association with carcinoma of the cervix. *Science* **161**: 1255, 1968.

(28) Terris, M.: The relationship of coitus to carcinoma of the cervix. *Amer. J. Pub. Health* **57**: 840, 1967.

(29) Martin, C. E.: Marital and coital factors in cervical cancer. *Amer. J. Pub. Health* **57**: 803, 1967.

(30) Stewart, H. L., Dunham, L. J., Caspar, J., Dorn, H. F., Thomas, L. B., Edgcomb, J. H. and Symeonidis, A.: Epidemiology of cancers of uterine cervix and corpus, breast and ovary in Israel and New York City. *J. Natl. Ca. Inst.* **37**: 1, 1966.

(31) Levitan, M. and Montagu A.: *A Textbook of Human Genetics.* New York: Oxford University Press, 1971.

(32) Pridan, H. and Lilienfeld, A. M.: Carcinoma of the cervix in Jewish women in Israel, 1960–67. *Israeli J. Med. Sci.* 7(2): 1465, 1971.

(33) Rotkin, I. D.: A comparison review of key epidemiological studies in cervical cancer related to current searches for transmissible agents. *Cancer Res.* **33**: 1353, 1973.

(received 7 April 1975)