PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Causality in medicine: the case of tumours and viruses

Vladimír Vonka

Phil. Trans. R. Soc. Lond. B 2000 **355**, 1831-1841
doi: 10.1098/rstb.2000.0738

THE ROYAL SOCIETY

doi 10.1098/rstb.2000.0738

Causality in medicine: the case of tumours and viruses

Vladimír Vonka

Department of Experimental Virology, Institute of Hematology and Blood Transfusion, Prague, Czech Republic (vonka@uhkt.cz)

Clarification of the aetiology of chronic human diseases such as atherosclerosis or cancer is one of the dominant topics in contemporary medical research. It is believed that identification of the causal factors will enable more efficient prevention and diagnosis of these diseases and, in some instances, also permit more effective therapy. The task is difficult because of the multistep and multifactorial origin of these diseases. A special case in contemporary aetiological studies is definition of the role of viruses in the pathogenesis of human cancer. Virus-associated cancer develops only in a small minority of infected subjects, which implies that, if the virus does play a role in the pathogenesis of the malignancy, other factors must also be involved. In this paper the author attempts to review the present methodological approaches to aetiological studies of chronic diseases, discusses the role of criteria for identifying causal relationships and proposes guidelines that might help to determine the role of viruses in human cancer.

Keywords: causality in medicine; viruses; cancer

'Conclusiveness in inferring causality is a desire more often than an accomplishment.'

(Susser 1988, p. 55)

1. INTRODUCTION

The problem of causality is one of the central themes of human thought. Paradoxically, philosophers have devoted much more attention to causality than have scientists, one of whose primary tasks it is to identify concrete causal relationships. In medicine, efforts to understand in greater depth the processes of recognition of causes of phenomena, and the logical and philosophical aspects of the problem, have only gathered in strength over the last three or four decades. This rise of interest in the subject has been associated with the contemporary state and tasks of medicine. Foremost among them is the challenge of the chronic maladies such as cardiovascular diseases and malignant tumours. In the developed countries they together account for 80% of human deaths. Identification of their causes is a basic precondition of getting them successfully under control. The problem is not at all simple. In either case we are facing pathological conditions whose development is multistep and multifactorial, i.e. complex and not easy to decipher from the causative point of view. This opens a wide gap for the most curious hypotheses, which, despite their poor substantiation, surprisingly often find their way as guidelines into medical practice and do not always guide it in a desirable direction. Among the medical community, the conviction has accordingly been gaining ground that definite and generally acceptable criteria should be set down, whose fulfilment could confirm or disprove an aetiological hypothesis and would entitle one to take practical steps towards prevention, diagnosis or therapy. An outstanding example and a still functioning catalyst of aetiological studies is the celebrated report of the USA Surgeon General on the relationship between smoking and lung cancer (US Surgeon General Advisory Committee 1964).

2. DEDUCTION, INDUCTION, KARL POPPER AND SCIENTIFIC HYPOTHESES

In medicine, the first signal of a possible causal relationship comes from the observation of repeated coincidence between factor F and disease D, and from the realization that their association in time and space is more frequent than accidental. Such an observation will give rise to the formulation of aetiological hypotheses. It is then a matter of investigation to decide between, or, more precisely, give preference to one of, the alternative hypotheses. There exist two basic logical procedures for drawing up causal conclusions: deduction and induction. They do not obey the same rules. By deduction we understand inference of particular propositions from general propositions. Deductive logic is a system of rules for getting true conclusions from true premises (Rothman 1988). Induction is the inverse process. It consists in inferring general rules, laws and predictions from repeated observations of phenomena. As Rothman (1988) points out, a major part of the philosophy of science has been dedicated, since the time David Hume subjected induction as a method of thought to a crushing criticism, to the revindication of induction or to attempts to conceive an epistemology of science without induction. The critics of induction claim that regardless of how carefully the inductive procedures are applied they do not warrant correct conclusions. This is so because they are immediately dependent on observations, which are not reliable enough owing to their limitations, finitude, and imperceptible connotations. Besides, there are no such things as completely impartial observations. Observations are probably always selective because they presuppose a certain task,

Received 16 February 2000 *Accepted* 8 May 2000

attitude, antecedent opinion, beliefs, a preconceived hypothesis, which—to a large extent—determine what the observer sees and what tends to be neglected because it is counter to the goal, i.e. they are subject to a tendency called 'wish bias' (Wynder *et al*. 1990).

Regarding scientific cognition and especially the inductive method, two camps have formed among the communities of philosophers and philosophizing scientists. Verificationists infer causal conclusions from repeated observations by means of the inductive method. Falsificationists, on the other hand, reject induction and abide by deduction alone, claiming that observations can only serve to reject theories.

The chief protagonists of the inductive method in the 20th century were Bertrand Russell and some members of the Vienna Circle (Gillies 1993). In an attempt to put inductive inference into a logical form Russell formulated the s.c. principle of induction (Russell 1998), comprising two statements. First, as Russell puts it, the greater the number of cases in which two things are associated, the greater the probability that they will be associated in a fresh case in which one of them is present. Second, a sufficient number of cases of associations and no cases of failure of such associations will make the probability of a fresh association nearly a certainty and will make it approach certainty without limit. The inconsistencies arising from the principle of induction have been criticized in the 20th century by many philosophers and scientists. Out of the critics of the induction method, two are most prominent, namely the French physicist, philosopher and historian of science Pierre Duhem and the recently deceased British philosopher of Austrian birth, Sir Karl Popper. Their criticisms differed in important respects. Duhem's approach can be labelled as historical. He criticized philosophical interpretations of concrete scientific discoveries, e.g. Newton's or Ampere's inductivism (Duhem 1991). On the other hand, Popper—being influenced by the Vienna Circle, of which he had never been a member and with the philosophical concepts of which he disagreed—adopted a strictly logical approach. In addition to his criticism of inductivism, Popper created a new general and consistent concept of philosophy of science called critical rationalism, which has markedly influenced the methodology of contemporary natural sciences (Popper 1959*a*,*b*,*c*,*d*, 1995). Let us take a look at the main principles of Popper's epistemology of science since they are relevant to our discussion.

Induction, says Popper (in agreement with Hume), is a psychological, not a logical process. Science is striding ahead leaning on deduction alone. The first step is the formation of a hypothesis, and this is an act of invention and imagination. Attempts to verify a hypothesis are devious and useless. Real progress is only possible when we endeavour to disprove the hypothesis. A useful hypothesis must therefore be testable. It should be simple, precise, and comprehensive, for this facilitates its testability. If attempts to falsify a hypothesis fail, it is neither verified nor refuted but remains a conjecture bound to be refuted or modified in the future. Thus, the best we can hope for is a theory that is nearer to the truth than its predecessors. As expressed in Jacobsen's schema (figure 1) (Jacobsen 1988), science in Popper's concept is a never-ending sequence of conjectures and refutations.

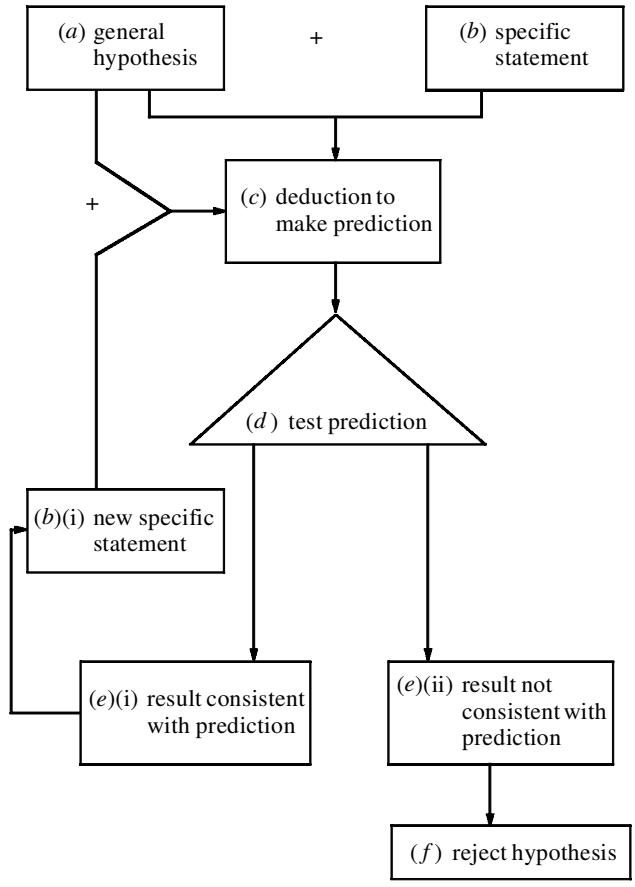

Figure 1. A logico-deductive (Popperian) model of scientific discovery (after Jacobsen 1988).

The present, Popper-inspired contention between verificationists and falsificationists is having repercussions of extraordinary intensity also in medicine, as heated discussions on the pages of journals and at various meetings testify.

Is there a possibility of reconciliation between the two trends? I daresay there is, and a reasonable bridging of the seemingly abysmal differences is, for medicine as a science standing in the direct service of man, indispensable. It is not the purpose of the following discussion to provide any definitive conclusions concerning the point, but rather to stimulate further relevant thought on the methodology of aetiological studies.

Our discussion will have five points of departure.

(i) Scientific hypotheses will be consistently distinguished from unscientific by their testability, i.e. only testable hypotheses can be considered scientific.

(ii) Falsification will be accepted as the critical element of scientific progress. However, this does not, in practical scientific work, preclude the simultaneous use of both falsification and verification, especially so if, together with Popper, we look at verification as failure of falsification attempts.

(iii) In our discussion we will not take up the position of militant deductivism. It appears to be ever more acceptable that new hypotheses in biomedicine are generated as a result of reflections on both preceding theory and new observations, i.e. of a certain interaction between deduction and induction. The efforts

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

of some of the contemporary philosophers to introduce terms like 'conjectural induction' or to differentiate between 'creative induction' and 'mechanical (or Baconian) induction' (Gillies 1993) are an expression of this understanding. Moreover, it is not the intention of this essay to contribute to the frequently unfair arguing of some of the present philosophers of science who seem to enjoy culling certain statements concerning induction or deduction out of context, in order to demonstrate striking logical inconsistencies in their opponents' concepts.

(iv) We will accept the role of the scientific community in assigning value to hypotheses. While I assume a wide consensus concerning acceptability of the first three points, I certainly do not take for granted positive attitudes towards the fourth. Rather, I anticipate the objection that I am adopting a non-scientific approach. In medicine, nevertheless, where any new conclusion concerning causality may have far-going consequences for human health, a different approach would be unjustifiable. The point is that search for causal relationships in medicine has, apart from epistemological and ontological aspects (which are inseparable), also its ethical side. In fact, the two processes, i.e. the process of drawing conclusions from aetiological studies and the process of making public health decisions (e.g. as to the introduction of a new vaccine), in spite of their different methodological concepts, strongly overlap and interact, rejecting the separation of theory and practice in medicine. If we take the consensus of the medical scientific community to be a necessary precondition for accepting a conclusion about something being a cause, we must do so with the reserve that this consensus does not warrant the correctness of the inference. Such consensus is always conditioned by the antecedent knowledge and its interpretation, and hence is time dependent. Thus, the consensus of the medico-scientific community can never have the weight of a final verdict and as such it should be associated with permanent criticism, which hopefully will induce corrective changes. Moreover, a consensus can be—and frequently is—influenced by a common lack of knowledge and also, unfortunately, by shared prejudices. These often stem from the paradigms of the time, which determine the manner in which scientists endeavour to explain unknown phenomena and form causal conclusions.

(v) Our discourse will in large part proceed along the pathways of epidemiology. This will be so not only because epidemiology, of all the medical sciences, probably stands closest to philosophy, but especially because the causes of those diseases that are the greatest concern of today's medicine will mostly be determined with the essential help of epidemiological methods. Looking at things from this angle, it is not surprising that most of the most ardent disputes about the ways of formation of scientific hypotheses in medicine and the methodology of their verification and/or falsification have been taking place among epidemiologists. Moreover, it is the conceptualization of causal models by which, in my view, the science of epidemiology has most significantly contributed to progress in theoretical medicine.

Our reasoning about the validity of theories will be kept predominantly within the limits of the contention between inductionists and deductionists. It should be noted, however, that the importance of other issues in connection with the verification or falsification, which do not completely rely on logical empiricism, has also been accentuated. These include relevance of the question of how a particular theory has been constructed, as reviewed and documented by examples from the physical sciences by Worrall (1985), and the so-called 'new experimentalism', which emphasizes the role of the actual experimental process and its analysis and interpretation in the inquiry into the inferences made (Hacking 1983; Mayo 1996*a*). These procedures are based on, and structured by, critically evaluated statistical tools (Mayo 1996*b*).

3. GENERAL SCHEME OF CAUSAL RELATIONSHIPS

Table 1 presents a general logical scheme of relationships between putative causative factor F (exposure) and disease D. In describing these relationships we shall use the philosophical categories of 'necessary' and 'sufficient' condition of cause. As pointed out by Susser (1973), there exist four types of possible relationships between F and D. F is 1, necessary and sufficient; 2, necessary but not sufficient; 3, not necessary but sufficient; or 4, not necessary and not sufficient. It is evident that there is an essential difference between patterns 1 and 2. Patterns 3 and 4 resemble them, but differ from them by conceding that a given clinical entity may have alternative causes. Though this is often so, our discussion will be confined to possibilities 1 and 2. The validity of our conclusions will thus be restricted to all cases of diseases that share a specific causative factor. Not only is the generation of any single disease that may have two or more alternative causes a much more complicated matter to consider, but there is also good reason to hold that such a clinical entity is not, *sensu stricto*, one disease but will in the future undergo subdivision according to its aetiologies. This will almost certainly be linked with redefinition of the descriptive criteria on the basis of which the disease is classified. On the other hand, we consider it highly expedient to discuss pattern 2 in two separate variants. Variant one: the causative factors act simultaneously—let us designate this subpattern 2A and say it represents a 'composite cause'. Variant two: the causative events come consecutively, and the subpattern 2B will be said to represent a 'catenated cause'. As will be shown below (see § 5) the methodology of investigating causal relationships differs for these two types of situation. Among the causative factors, events which create conditions for the cofactors to occur also have to be included. In the case of cancer they are called 'promotors', which are defined as events that are not carcinogenic by themselves but help the development of cancer.

It should be mentioned that the causal factors additional to F (whether in a composite or catenated cause, i.e. pattern 2A or 2B) must be differentiated from the events mediating the transmission from F to D.

Table 1. *General logical patterns of relationship between causative factor (F) and disease (D) (modified after Susser 1973)*

pattern	F is necessary	F is sufficient
1	+	+
2	+	−
3	−	+
4	−	−

1. F is necessary and sufficient to produce D
   $F \rightarrow D$
   (i.e. if F is present, non-D is impossible).

2. F is necessary but not sufficient to produce D

   A. Cofactors Cf_1–Cf_n act simultaneously with F

   F + Cf_1 . . . Cf_n → D

   B. Cofactors Cf_1–Cf_n act consecutively after F
   $F + Cf_1–Cf_n \rightarrow D$

3. F is not necessary but is sufficient to produce D
   $F \rightarrow D$
   $F_1 \rightarrow D$
   $F_2 \rightarrow D$

4. F is neither necessary nor sufficient to produce D
   $F + Cf_1 \ldots Cf_n \rightarrow D$
   $F_1 + Cf_1 \ldots Cf_n \rightarrow D$
   $F_2 + Cf_1 \ldots Cf_n \rightarrow D$

Discontinuity is one of the basic attributes of the cause–effect relationship in biomedicine. Between F and D there is usually a chain of events that are connected to F at one end and to D at the other end. Their number is dependent on the nature of the process but also on the level at which these events are defined. In our search of causes we will consciously omit this chain of events which mediates the F→D sequence and shall consider it a standard condition.

4. CRITERIA FOR IDENTIFYING CAUSAL RELATIONSHIPS

In addition to the discontinuity just mentioned, the relationship between cause and effect in medicine has three basic characteristics: association (i.e. coincidence of F and D, see § 3); temporal sequence, i.e. the cause is antecedent to the effect; and direction, i.e. F leads to D. Association is evaluated in relation to chance concurrence. It has to be demonstrated by statistical methods as falling outside the limits of expected variation and persisting under variable conditions. It is evident that in biological systems the cause is always antecedent to the effect, but it is not always easy to demonstrate this; indeed, sometimes it is very difficult to do so (see below, this section). Direction is an expression of asymmetry between cause and effect (by definition, symmetrical associations are non-causal). This asymmetry is the point of departure in compiling sets of criteria proposed and used for clarifying causal relationships. According to many authors—and these are not only outright verificationists—the fulfilment of such a set of criteria affirms or makes highly probable the existence of a postulated causal relationship. We will be a bit more modest and speak of corroboration rather than confirmation.

Table 2. *The Hill criteria for drawing causal conclusions (Hill 1965)*

criterion
1. strength of association
2. consistency
3. specificity
4. temporality
5. biological gradient
6. plausibility
7. coherence
8. experimental evidence
9. analogy

The most frequently discussed and the most generally accepted are the criteria formulated by Sir A. B. Hill (Hill 1965, p. 299) (table 2). It should be stressed that their author did not consider them as 'hard-and-fast rules of evidence that must be obeyed before we accept cause and effect'. Rather, he was convinced that none of his 'nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a *sine qua non*' and that they should function more as a stimulus 'to make up our minds on the fundamental question—is there any other way of explaining the set of facts before us; is there any other answer equally, or more, likely than cause and effect?' Even superficial examination of Hill's criteria reveals several characteristic points. First, some of the criteria overlap and complement each other (e.g. strength of association and specificity, or coherence and biological gradient) rather than being mutually disparate. Second, it is evident that their number does not need to be definitive. Some other criteria could be added or, as advocated by others, some of those listed could be excluded. Third, the criteria do not constitute any clear hierarchy, as one can easily imagine situations where some are dramatically more important than others, in other words, the value of individual criteria depends on the hypothesis tested. Fourth, since establishment of causal inference is an evolving process, the virtues of the various criteria are dependent on the stage of this process. Fifth, some of the criteria are more useful for corroborating and others for falsifying a hypothesis. Precisely this point makes it possible to judge their use from the point of view of the controversy between verificationists and falsificationists. However, since words may have different meanings, we must first be clear about what values the individual criteria signify. The present concepts of the criteria are somewhat different from Hill's original descriptions, which is a reflection of the developments over the last three decades. One of the most comprehensive recent discussions of these criteria is that by Susser (1988). The

following outline is partially based on his characterizations.

Strength of association is not equivalent to statistical significance. An association may be weak but at the same time—given a big enough population sample—be statistically highly significant. Strength of association is expressed as relative risk (RR) or the so-called odds ratio (OR). The higher the RR or OR, the higher the strength of association and the more probable the causal junction. Even very high strength of association cannot confirm a hypothesis. However, it may make it more preferable to other hypotheses. Although lack of significance does not provide a logical argument for the falsification of a hypothesis, statistics does play a key role in quantifying the strength of association, in discriminating between real effects and artefacts and, most importantly, in falsifying the hypothesis, provided that the study has been well designed and executed, the power has been adequate and the measurement error low.

Consistency means identity of findings in repeated studies carried out by different researchers, in different places and at different times. It would be a mistake to expect that the RR or OR values would be equivalent in them, however. The consistency criterion is qualitative, not quantitative. No event can be repeated precisely, because changes occur with time and the populations studied, and also with the design and execution of the study. Reproducibility is sometimes said to be an ideal representation of classical induction and is taken as the most convincing tool of verification by inductivists. Irreproducibility has the significance of falsification, provided that the particular study was well conceived and rightly executed and analysed.

Specificity of association is the degree of regularity with which the occurrence of F is connected with the occurrence of D. The closer the ratio approaches one, the more specific the relationship, and the more probable the association. The ideal ratio is 1:1, but it is almost never encountered in medicine. However, even very low specificity or its absence does not preclude a causal relationship and has no falsifying significance. The evidence that polioviruses cause poliomyelitis is not weakened by the infection taking a clinically inapparent course in a great majority of infected subjects or by its being more often manifested by a non-specific illness resembling influenza or aseptic meningitis than paralytic disease. It is a rule that specificity in cause is of greater importance than specificity in effect, because it suggests the absence of other causes. In general, the criterion of specificity should be treated cautiously. It should be recalled that the lack of specificity in the smoking–lung cancer relationship has been misused as an argument, raising doubts about the nature of this relationship.

Temporality, i.e. the time sequence, does not need special explanation. In spite of the logical clarity and demonstrational significance of this criterion, in chronic diseases it is frequently difficult to obtain evidence of its being fulfilled, however. It is a basic characteristic of retrospective (so-called case control) studies—which make up an overwhelming majority of aetiological investigations in medicine—that both the cause and the effect occurred in the past. Reconstruction of the sequence of events is furthermore impeded by the fact that most chronic diseases have an insidious onset and one cannot say exactly when a case started. Time sequence can only be determined reliably by prospective studies, in which the observer can step in between the cause and its effect. If an event precedes an effect, this is not evidence of its causative function, but it may supply a reason for investigating the relationship further. If, however, a reverse sequence is found, the hypothesis is thereby refuted.

The criteria of biological gradient, plausibility and coherence overlap strongly. Biological gradient means a proportionality of the response to the dose (at least within certain limits) or to the length of exposure. If observed, it has a corroborative effect. However, this plainly cannot be required in some situations, especially where threshold effects are frequently involved. Thus, absence of proportionality does not necessarily have a falsifying effect. Coherence is the measure of consistency with current biological knowledge. Speaking of the coherence of a hypothesis we have in mind the extent of its agreement with theoretical deductions or inductive conclusions drawn from previous observations. Sometimes a distinction is made between biological and epidemiological coherence, the latter being then usually called 'plausibility' and the term 'coherence' restricted to biological credibility.

Experimental evidence can be obtained in a number of ways. The most valuable is the demonstration that the effect will not come if the presumed cause is removed. Its principle has been neatly formulated by the famous 19th century French physiologist Claude Bernard in his *Introduction a l'étude de la medicine experimental*: 'The only proof that one phenomenon is the cause of another is that by removing the first we stop the second' (Bernard 1961). (It may be of interest that this is a paraphrase of a much older statement (by 600 years) made by the learned Dominican Thomas Aquinas: '*Sublata causa, tollitur effectus*'. It is also worthy of note that in Bernard's text the sentence cited is directly preceded by an encouragement that the scientist should try to disprove his hypothesis, which sounds very Popper-like indeed.) This principle, so brilliantly articulated by Bernard, is reflected by a trend in modern philosophy of science, which stresses the role of intervention in achieving the aim of scientific inquiry (Hacking 1983). The experimental evidence desired may come, for example, from intervention by a specific vaccine that will eliminate or lower the incidence of a certain disease. However, one should be careful when evaluating such data. The failure of an experimental vaccine that should, according to the underlying hypothesis, have reduced the incidence of a disease, needs not mean its refutation. The vaccine as it had been prepared, or when administered in the manner used, may not have been effective enough, or the type of immunity it induced may not have been able to prevent the pathogenic process.

Analogy probably does not need any explanation. An example of analogy: since it has been found that animal retroviruses can induce malignant growth in their natural hosts, it has become easier to assume that there exist human retroviruses that might give rise to tumours in man. From the point of view of deductive logic analogy is of little value. It is, however, a frequent source of hypotheses.

5. FROM CRITERIA TO AETIOLOGICAL HYPOTHESES

Let us now take a look at the power of Hill's individual criteria to falsify aetiological hypotheses.

The subsequent discussion and conclusions are based on the reasoning of Weed (1988). According to him, the possibility of directly deducing Hill's criteria from a causal hypothesis is largely dependent upon the general nature of the hypothesis. Thus, in the case of the necessary-and-sufficient cause hypothesis (table 1, pattern 1) and in the case of the composite cause hypothesis (pattern 2A) the possibility of using deductive reasoning is more limited than in the case of the catenated cause (pattern 2B). Judging by what we know of the pathogenesis of malignant tumours and atherosclerosis today, the development of these diseases corresponds very well with pattern 2B. A multistage process that involves causes acting one after the other permits more precise and more detailed deductions and hence more facile falsifications. Direct deductions are possible with the strength of association, consistency, temporality, experimental evidence and biological gradient. Problems remain with specificity, coherence and analogy. They are not deducible from any of the hypotheses discussed. However, while both coherence and specificity assume the act of deduction, analogy has little to do with deductive logic. This critical evaluation is of importance for making practical decisions about causality; the criteria can mainly contribute if they falsify the causal hypotheses tested or competing explanations.

What, then, is the Popperian alternative to the Hill (or other sets of) criteria? Is it refusal of whatever criteria on account of their being, as radical Popperians maintain, mere instruments for confirming beliefs? D. L. Weed, a more moderate adherent of Popper, proposes that all criteria be compacted into two, namely predictability, i.e. the ability of a hypothesis to foretell unknown events, and testability (Weed 1988). From any causal hypothesis, regardless of its kind, predictions can be derived. (In this case, as pointed out by Worrall (1985), a strict distinction should be made between the results used in the construction of the hypothesis and those predicted by the hypothesis, i.e. the predictive value could only be applied to situations not included among those on which the hypothesis has been constructed.) Yet, this will not do; another condition must be fulfilled—what has been predicted must be testable. Thus, testability links prediction with observations. Causal hypotheses should accordingly be so formulated that their conjunction with specific statements may enable predictions to be deduced and then tested. The higher the precision of the prediction, the more probable the possibility of its conflict with the subsequent observation. If the outcome does not agree with the prediction, the hypothesis will be rejected. On the other hand, the more such tests it survives, the greater its contribution to knowledge, the closer its approximation to truth, the higher its preference over other hypotheses.

But what of analogy? Is it possible or legitimate to condemn it outright? From the Popperian angle, it has no place in falsification of causal hypotheses. Nevertheless, the medico-biological sciences cannot reject it as an auxiliary tool. Analogy is used in the generation of hypotheses, and to expel it dogmatically from the process of cognition would be to give up an important helper in this respect. One can, however, go so far as to agree with Popperians that it is not a very ingenious source of hypotheses and that creative invention is more productive and of greater intellectual consequence.

6. VIRUSES AND TUMOURS: BASIC CHARACTERISTICS OF THE SYSTEM

After a somewhat lengthy but necessary and hopefully sufficing preparatory discourse, let us turn to the subject proper of this essay, which is a discussion of the role of viruses in the pathogenesis of malignant tumours in man. According to what we know today, viruses are directly involved in the development of about 15% of human tumours. (We will leave out of our discussion situations in which viruses contribute to tumour development indirectly by inducing a state of immunosuppression, e.g. in AIDS. This falls under a broader category of pathological phenomena resulting from weakened efficacy of immune surveillance.)

To demonstrate that one or another virus is an agent of a human tumour is a difficult matter—indeed, it probably belongs among the most difficult tasks of biomedical research. The problem is complicated primarily by the fact that all of the incriminated tumour viruses are widely distributed among human populations and, in so far as they give rise to tumours, they do so in only a low percentage of infected individuals. Most of the infections are clinically inapparent and so specificity of the effect is very low. It is thus clear that a virus is not a sufficient cause of the disease. This conclusion corresponds well with the contemporary knowledge of the biology of normal cell transformation into a tumour cell. A tumour cell does not arise all at once but is a result of gradual development on a background of alterations of three categories of its genes. The first two are known as 'oncogenes' and 'tumour suppressor genes', and their products, respectively, positively and negatively influence cell multiplication and survival, while the third category consists of genes whose products take part in cellular DNA repair. The products of tumour virus genes intervene significantly in cell growth control, but under natural conditions are apparently unable to affect the development of a tumour cell by themselves. Insertion of the viral genetic material into the cell (though not necessarily integrated into the host cell chromosomes) and expression of the viral oncogenes are events that must be followed by other events on the way to malignant transformation of the cell. There is also another mechanism of tumour cell induction by viruses, namely insertional mutagenesis occasioned by integration of viral DNA (DNA copy of viral RNA) into the host cell genome. Even this event does not suffice, it definitely has to be followed by other genetic alterations. In terms of the logical patterns of causal relationships (see table 1), virus-associated malignant transformation is a typical instance of pattern 2B, which corresponds to the general hypothesis of catenated cause.

7. EVOLUTION OF CAUSAL THOUGHT IN MICROBIOLOGY

Before we attempt to systematize the ways and means appropriate for critical evaluation of causal relationships

between viruses and tumours, it will be useful to summarize briefly the development of causal thinking in microbiology.

The problems of causal relationship between microbe and disease came to the forefront of interest for the first time towards the end of the 19th century when, following the introduction of solid media in bacteriology, newly isolated bacterial species began to become known quickly. 'Conclusions' concerning the role of the different agents in the pathogenesis of human and animal diseases began accumulating equally fast. The first scientist to realize that it was necessary to introduce order into the interpretation of findings was the celebrated German bacteriologist Robert Koch (Koch 1892). More than a hundred years ago he formulated principles, known as Koch's postulates, by whose fulfilment he conditioned confirmation of causal relationship between a microbe and a disease. Their briefest version is as follows. The particular infectious agent: (i) is present in every case of the disease, under conditions that may account for the pathological changes and clinical course of the illness; (ii) is not present in other diseases as an accidental and non-pathogenic agent; (iii) after it has been isolated and grown in pure culture, is capable of producing the disease.

As is evident, the criteria are very rigorous. Their formulation concedes nothing but sufficient causation (pattern 1 in table 1) and maximum specificity and do not allow for the possibility that pathogenicity of a microbe is strictly species specific. The first two postulates are inductive, the third is deductive. In Koch's conception, a negative result was tantamount to falsification.

Although Koch's postulates have done a good service to medicine by helping to clarify the aetiology of many infectious diseases and influencing medical thought positively even beyond the area of infectious diseases, their uncompromising application has had some adverse affects as well. Sometimes it contributed to darkening, rather than throwing light on, the origin of a disease. With the present state of knowledge we can easily see all that Koch's postulates left out of view: the asymptomatic-carrier status, epidemiological and immunological aspects of the study of causality, existence and reactivation of latent infection, autoimmune processes, etc. Koch himself later realized the excessive rigidity of the principles he had set up. Nevertheless, I am not aware he ever tried to reformulate them.

Notwithstanding all the objections we may have against them, Koch's postulates have, with various modifications, survived until these days. Their first important complementation was undertaken in the 1930s, during the first period of modern virology, when serological methods began to have their say. The modification introduced by the American virologist T. Rivers amounted to recognition of specific antibodies, namely their appearance or rise of their level, as evidence equivalent in significance to isolation of virus (Rivers 1937). This helped medical virology enormously. Let us not forget that until the beginning of the 1950s virus isolation was a very exacting accomplishment and the prerogative of a limited number of laboratories. A turning point came in the 1950s after the massive introduction of tissue cultures into medical virology that fundamentally simplified and facilitated the isolation of viruses. In the course of only a few years thousands of new virus isolates representing over 100 new virus species were heaped up in virology laboratories. Most often they came from patients suffering from undifferentiated diseases of the upper respiratory tract or the alimentary tract, or from aseptic meningitis. However, the same viruses were very often also found in healthy subjects, which suggested that, in so far as they caused illness at all, they only did so in rare cases. Moreover, no susceptible host was found among laboratory animals for a large majority of the new viruses. Based on these findings, new guidelines for determining aetiological relationships were formulated by Huebner (1957). They were probably the most important modification of Koch's postulates. The most essential change proposed by Huebner amounted to the new postulate of conditioning the proof of causal relationship between a virus and a disease by epidemiological evidence. This amounted to demonstration that the virus was present significantly more often in patients than normal individuals of the same age and sex, living in the same place at the same time, and under comparable socio-economic conditions. Indeed, epidemiological investigations turned out to be the messenger that brought convincing evidence that many of the new viruses were pathogenic for humans.

The changes in thinking brought about by the new findings and the concepts that germinated upon them left their mark on the general understanding of the pathogenesis of infectious diseases. Since the late 1950s there has been a steady shift of interest from the infectious agent alone to the circumstances of the infection and the response of the host. In other words, it has transpired that the development of illness does not depend on the presence of the infectious agent alone, but also on individual predispositions, i.e. on the physiological state and genetically conditioned susceptibility of the infected individual, as well as on participation of external co-factors that may lower his resistance. This means a definitive departure from the hypothesis of the sufficient cause resting with microbiology.

Table 3. *Criteria of aetiological relationship between virus and tumour (Evans 1982)*

immunological
presence of virus antibodies is more frequent in patients than in control subjects
raised antibody levels are more frequent in patients than in control subjects
virus antibodies and their raised levels precede development of tumour
virological
virus genetic material is present in tumour tissue but absent from normal tissues
the virus is capable of transforming normal cells into cells carrying tumour potential
experimental
the virus is capable of producing tumours in susceptible animals
in experimental tumours, the virus or its genetic material can be detected
neutralization of the virus prior to inoculation will prevent tumour development

Table 4. *Criteria of aetiological relationship between virus and tumour (Zur Hausen 1991)*

regular presence and persistence of nucleic acid of the virus or of a related type in cells of specific malignant tumours
induction of proliferative changes upon transfection of the virus genome or parts therefrom in appropriate tissue culture cells
demonstration that the induction of these changes and the malignant phenotype of the tumour cells depend on functions exerted by the persisting viral DNA
epidemiological evidence that infections with the respective virus represent risk factors for tumour development

Table 5. *Tests for hypotheses of aetiological relationship between virus and tumour*

direct tests
epidemiological
testing for prevalence of infection by the virus in patients and control subjects—determination of strength and consistency of this association
testing for epidemiological characteristics of tumour occurrence and spread of virus
testing for time sequence
testing for effect of intervention against virus on tumour incidence
immunological
testing for immune reactivity of patients with virus antigens, especially those present in tumour cells
testing for relationship between strength and character of specific immune responses and clinical state
molecular–biological
testing for the presence of virus-specific macromolecules in tumour cells
indirect tests
testing for oncogenicity of virus in laboratory animals
testing for capability of virus to transform cells cultivated *in vitro*
testing for persistence of virus in infected organism, etc.

8. METHODOLOGY OF THE STUDY OF VIRUS–TUMOUR RELATIONSHIPS

Modern tumour virology was born in the 1960s. Gradually, a long series of viruses eliciting tumours in animals were identified, and initially indirect and then direct evidence that viruses were the agents of some forms of malignant growth in humans began to accumulate. The first attempts to introduce order into judgement on aetiological relationships started from Koch's postulates. It soon became clear, however, that not one of them could be adhered to in the study of the role of viruses in the development of human tumours. This conclusion followed not only from the above-mentioned extensive distribution of potential tumour viruses in human populations and the high frequency of asymptomatic infections, but also from the gradual recognition of the nature of the virus–cell interaction that leads to malignant transformation. In short, it became clear that there were no scientific grounds for the postulate of presence of infectious virus in the tumour or its absence in healthy subjects or, considering the frequent species restriction of the oncogenic effect, for the strict postulate of reproducibility of the disease in laboratory animals.

Accordingly, the need was realized for the formulation of a new set of postulates, criteria or guidelines the application of which would enable more accurate judgement of the role of viruses in the development of malignant tumours in man. The progress made in the methodology of aetiological studies of non-infectious diseases played a significant part in their elaboration. Stimuli came from the above-mentioned demonstration of causal relationship between smoking and lung cancer as well as from the discovery of an association between diet and cardiovascular diseases. It became furthermore increasingly evident that aetiological conclusions in chronic diseases, including malignant tumours, must be based on integrated findings of different medical disciplines, among which epidemiology was to play a prime role, and not on any separate finding, no matter how significant. It was at this time that Hill's criteria were drawn up, to be either preceded or followed by sets compiled by Yerushalmy & Palmer (1959), Lilienfeld (1959), Sartwell (1960), Susser (1977), Johnson & Gibbs (1974) and some others.

The nature of the problem in tumour virology was neatly expounded by the late Alfred S. Evans, who paid systematic attention to causality issues in medicine. In his considerations, he put special emphasis on the historical background of the various solutions, which stemmed up not only from the growing volume of knowledge, but also from the evolving medico-biological technologies. He attempted to surmount the existing conceptual limitations and the unhappy schism that arose and gradually intensified in discussions on the aetiology of acute and chronic diseases. He proposed a new set of criteria for clarifying the relationships between viruses and tumours (Evans 1982), presented in table 3. As may be seen, Evans divided his criteria into three categories, namely immunological (in fact, epidemiological), virological and experimental, and was actually the first in tumour virology to respond in a creative way to the demand for interdisciplinarity. Each category includes a number of partial postulates whose fulfilment conditions causal inference. In formulating his criteria Evans was strongly impressed by the growing evidence that the Epstein–Barr virus was the causative agent of Burkitt's lymphoma and nasopharyngeal carcinoma, and that another of the herpesviruses, the agent of genital herpes (so-called herpes simplex virus type 2, HSV2) was aetiologically involved in cervical carcinoma.

However, the situation changed markedly in subsequent years, partly owing to falsification of the hypothesis of the causal role of HSV2 in the development of cervical carcinoma and partly owing to discoveries of new human tumour virus candidates, i.e. the viruses of hepatitis B and C, papillomaviruses, and human T leukaemia virus types I and II (HTLV-I and HTLV-II). These findings could only with difficulty be compressed into Evans' schema as drawn up in the early 1980s. This was recognized by H. zur Hausen, the pioneer in investigating the role of papillomavirus in cervical cancer. Among the criteria he especially required the regular demonstration of viral DNA in tumour cells and the proof that the malignant phenotype of transformed cells depends on it. He formulated four postulates 'which, if fulfilled, should permit an unambiguous identification of human or animal tumour viruses' (Zur Hausen 1991,

Table 6. *Significance of epidemiological findings for verification or falsification of hypotheses*

(Single cross, significance low; double cross, significance high.)

result	characteristics	verification	falsification
prevalence of infection higher in patients than controls	strength of association		
	high	++	
	low	+	
	none		+
	consistency		
	yes	++	
	no		+
epidemiological features of tumour occurrence and virus spread are similar	yes	+	
	no		+
tumour preceded by infection	yes	+	
	no		++
incidence lowered by intervention against virus	yes	++	
	no		+

p. 687). Zur Hausen's postulates are summarized in table 4.

Already before that, we also ventured to modify Evans' postulates (Vonka *et al.* 1987). The set of criteria we suggested for the most part represented a combination of the aforementioned elements and concepts. The main innovations were (i) an emphasized priority of epidemiological criteria, (ii) a certain mitigation of the significance of the virological findings themselves, and (iii) categorization of the different elements of evidence into the more important direct (i.e. decisive) and the less important indirect (i.e. supportive). Although, in my opinion, no important new findings that would necessitate a fundamental factual revision of these guidelines have appeared since their publication, in the light of the dazzling development of the philosophy of science and its ever increasing penetration into the sphere of biomedical research, they do not seem acceptable in the form originally presented. I have accordingly tried to transform them in line with Popper's epistemology, replacing the terms 'criterion' and 'postulate' (terms with which many people find fault) by formulations implying that the essence of aetiological studies is testing the hypotheses and predictions inferred. Reformulated, they are presented in table 5.

Let us now examine the capacity of the different tests to verify or falsify a hypothesis, in dependence on the character of their results. Table 6 attempts to evaluate the cognitive power of epidemiological studies. It is clear that some of the results will have a verifying and others a falsifying meaning. Let us not be deceived by the almost even distribution of the crosses denoting the significance of different results. In order to have verifying value positive findings would have to be obtained in all or nearly all tests, whereas falsification in any one point will falsify the hypothesis as a whole. We (Vonka *et al.* 1984) and later other groups (Adam *et al.* 1985; Lehtinen *et al.* 1993) have falsified the hypothesis of an aetiological role of HSV2 in the pathogenesis of cervical carcinoma in prospective studies by the sole demonstration that there is no excess risk of cervical neoplasia in HSV2-infected women, thus indicating that development of the neoplasia is not preceded by HSV2 infection. Perhaps this example could illustrate the asymmetry between falsification and verification so strongly accentuated by Popper.

A similar picture is presented by the results of direct molecular biological and immunological tests (table 7). Their results seem to be of lesser value both in terms of verification and falsification. Inferior still is the value of results of indirect tests, whose inclusion in the scheme respects the principle of analogy (table 8). They can certainly play the role of a corroborative factor in verification attempts but have no falsification value.

There are two caveats and provisos. First, the demonstration of markers of viral infection (antibody, viral nucleic acid) in patients and in a control population may not be reliable enough because of the lack of adequate laboratory technology. This was an inherent shortcoming of studies on the HSV2–cervical cancer relationship in the 1960s and 1970s and papillomavirus–cervical cancer relationship in the 1980s. Thus, to be able to verify or falsify an aetiological hypothesis the measurement error should be very low. Even a low level of misclassification can considerably distort the expected prevalence of virus infections and induce incoherencies in epidemiological findings. Empirical simulations have shown that low specificity represents a much greater problem than low sensitivity (Franco 1991). Second, some evidence has accumulated suggesting that some viruses could transform cells in a 'hit-and-run' manner, like chemical or physical carcinogens (Schlehofer & Zur Hausen 1982). This would have an impact on the methodology proposed. Out of the three categories of criteria listed in the preceding tables, only the epidemiological ones could be applied, and epidemiology alone, then, should provide the evidence of involvement of a virus in cancer.

9. CONCLUSIONS

Let us return from tumour viruses to the general field of discourse. I daresay we should approach the current philosophical disputes on the nature of causality in science with a large dose of tolerance but not reject the lessons that modern epistemology offers. Although we do not need to accept all of Popper's concepts, we should appropriately appreciate his contribution to the endeavour to find and define in a new way the logical relationship between observation and scientific theory. The methodology based on conjecture and rejection opens

Table 7. *Significance of immunological and molecular biological findings for verification and falsification of hypothesis*

result	characteristic	verification	falsification
patients react with virus-specific antigens of tumour cells	always	++	
	sometimes		
	never		+[a]
relationship between level and nature of specific immune response(s) and clinical state in patients present	always	++	
	sometimes		
	never		+[a]
virus-specific macromolecules present in tumour cells	always	++	
	sometimes		
	never		+[a]

[a] There is a theoretical possibility that in some situations the virus acts like a chemical carcinogen (in a 'hit-and-run' manner); hence a negative result need not have a strongly falsifying meaning.

Table 8. *Significance of indirect tests for verification and falsification of hypotheses*

result	characteristic	verification	falsification
virus produces tumours in animals	yes	+	
	no		
virus transforms cells *in vitro*	yes	+	
	no		
virus persists in infected organism for prolonged periods of time	yes	+	
	no		

space for, and invites, critical thinking, encourages those coming with new ideas, frees cognition of its subjective dimension, and warns us against mistaking presumed knowledge (belief in knowledge) for real knowledge. It restores to the process of cognition its empirical basis and supplies evidence that cognition grows without relying on verification. The legitimacy of Popper's conviction, i.e. soulless piling up of 'corroborative' observations is of far lesser value for scientific progress than is active effort to overthrow hypotheses and replace them by others, could be demonstrated in many instances.

However, nothing was more alien to the great thinker and his philosophy than doctrinarianism, and a mistake is committed by some extremists among his adherents who would like to topple down some of the principles on which research was based in the medical sciences over the past 100–150 years. Their rejection of whatever order or frame in scientific work is not only unjust, factually incorrect and scientifically uneconomical, but in a way is also risky in the context of medical investigation. In medicine, the problem of causality has a specific feature that follows, as already pointed out, from possibly very serious consequences of any new causal conclusions. For the evaluation of their validity and for their consequent possible acceptance or not by the medical community, some set of rules is needed which, notwithstanding its limitations, furnishes the researcher with guidelines as to what questions should be addressed and what answers should be sought, and, furthermore, inspires the scientific community towards critical evaluation of the findings obtained. These rules, if thoughtfully applied, should improve the scientists' skills in solving their problems. Although general consensus about any such set of rules assisting judgement about causality cannot even be expected, at least a wide agreement following from critical evaluation of the various sets proposed is extremely desirable. Such a set or system of guidelines, if it is to be functional, should institutionalize as a working concept an interdisciplinary approach to the investigation of causal relationships between a putative aetiological factor and a disease under study, which I have attempted to illustrate with the example of viruses and tumours.

Whatever the system of guidelines, however, it must never be a strait-jacket limiting the freedom of research. Attempts at a strict application of philosophical principles in science would necessarily separate the logic of science from the practice of science, which would bring about a most undesirable state. As Susser (1988) observes, if a research worker must decide what to do next 'in the narrows and rapids of research', then he will prefer to follow his common sense rather than philosophical precepts. Even if falsification is of much higher significance than verification in the process of recognition of causes, without verification—which can be defined as a high degree of probability not readily open to doubt because of not being falsifiable by contemporary means—we could probably not work our way to any conclusions acceptable to the medical community and usable in medical practice. However, such conclusions should be drawn with the utmost care, with humility before the yet unknown. Even if medical conscience tends to oppose the thesis that causal relationships cannot be proved beyond any doubt, one must face it. Only in this way we can try to avoid a greater number of blunders than the unavoidable. The most practical concrete consequence of such a stance is a high degree of circumspection in the introduction of

new therapeutic, diagnostic and preventive procedures. From the epistemological point of view, such an approach warrants continuation of the process of cognition, and is essentially, whether we like it or not, a further invitation to falsification.

The author wishes to thank his many friends and colleagues in Canada, the Czech Republic, Germany, Sweden, the United Kingdom and the United States for very useful discussions and their comments on the manuscript in the course of preparation of this paper. The research work of the author is supported by grant no. 312/99/0542 of the Granting Agency of the Czech Republic and grant no. 5959-3 of the Internal Granting Agency of Ministry of Health.

REFERENCES

Adam, E., Kaufman, R. H., Adler-Storthz, K., Melnick, J. L., Dreesman, G. R. A. 1985 Prospective study of association of herpes simplex virus and human papillomavirus infection with cervical neoplasia in women exposed to diethylstilbestrol *in utero*. *Int. J. Cancer* **35**, 19–26.

Bernard, C. 1961 *Introduction a l'etude de la medicine experimental* (ed. O. Poupa), pp. 115–136. Prague, Czech Republic: SZN Publishers. (Czech translation, originally published in 1865.)

Duhem, P. (ed.) 1991 Physical theory and experiment. In *The aim and structure of physical theory*, pp. 180–211. Princeton University Press. (English translation, originally published in 1904–1905.)

Evans, A. S. 1982 Viruses. In *Cancer epidemiology and prevention* (ed. D. Schottenfeld & J. Fraumeni), pp. 364–390. Philadelphia: W. B. Saunders Co.

Franco, E. L. 1991 The sexually transmitted disease model for cervical cancer: incoherent epidemiological findings and the role of misclassification of human papillomavirus infection. *Epidemiology* **2**, 98–106.

Gillies, D. (ed.) 1993 Some historical background: inductivism, Russell and the Cambridge School, the Vienna Circle and Popper. In *Philosophy of science in the twentieth century*, pp. 3–25. Oxford, UK: Blackwell Publishers.

Hacking, I. (ed.) 1983 Intervening. In *Representing and intervening: introductory topics in the philosophy of natural science*, pp. 149–275. Cambridge University Press.

Hill, A. B. 1965 The environment and disease: association or causation? *Proc. R. Soc. Med.* **58**, 295–300.

Huebner, R. J. 1957 The virologist's dilemma. *Annls NY Acad. Sci.* **67**, 430–442.

Jacobsen, M. 1988 Inference in epidemiology. In *Causal inference* (ed. K. J. Rothman), pp. 105–118. Chestnut Hill, MA: ERI.

Johnson, R. T. & Gibbs, C. J. 1974 Koch's postulates and slow infections of the nervous system. *Arch. Neurol.* **30**, 36–38.

Koch, R. 1892 Über bacteriologische Forschung. In *Vehrhandlungen des X. Internationalen Medizinisches Congresses Berlin, 4–9 August 1890*, pp. 35–47. Berlin: August Hirschwald Verlag.

Lehtinen, M., Hakama, M., Aaran, R.-K., Aromaa, A., Knekt, P., Leinikki, P., Maatela, J., Peto, R. & Teppo, L. 1993 Herpes simplex virus type 2 infection and cervical cancer: a prospective study of 12 years follow-up in Finland. *Cancer Causes Control* **3**, 333–338.

Lilienfeld, A. M. 1959 On the methodology of investigations of etiologic factors in chronic diseases—some comments. *J. Chronic Dis.* **10**, 41–46.

Mayo, D. G. (ed.) 1996*a* The new experimentalism and the Bayesian Way. In *Error and the growth of experimental knowledge*, pp. 57–101. Chicago and London: University of Chicago Press.

Mayo, D. G. (ed.) 1996*b* Why Pearson rejected the Neyman–Pearson (behavioristic) philosophy and a note on objectivity in statistics. In *Error and the growth of experimental knowledge*, pp. 361–411. Chicago and London: University of Chicago Press.

Popper, K. (ed.) 1959*a* A survey of some fundamental problems. In *The logic of scientific discovery*, pp. 27–47. New York: Basic Books, Inc. (Originally published in 1934.)

Popper, K. (ed.) 1959*b* Theories. In *The logic of scientific discovery*, pp. 59–77. New York: Basic Books, Inc. (Originally published in 1934.)

Popper, K. (ed.) 1959*c* Falsifiability. In *The logic of scientific discovery*, pp. 78–92. New York: Basic Books, Inc. (Originally published in 1934.)

Popper, K. (ed.) 1959*d* Corroboration, or how a theory stands up to tests. In *The logic of scientific discovery*, pp. 251–281. New York: Basic Books, Inc. (Originally published in 1934.)

Popper, K. (ed.) 1995 Induction, deduction, objective truth. In *Unended quest. An intellectual biography* (ed. M. Tarnovsky), pp. 135–142. Prague, Czech Republic: Prostor Publishers. (Czech Translation, originally published in 1976.)

Rivers, T. M. 1937 Viruses and Koch's postulates. *J. Bacteriol.* **33**, 1–12.

Rothman, K. J. (ed.) 1988 Inferring causal connections—habit, faith or logic? In *Causal inference*, pp. 2–12. Chestnut Hill, MA: ERI.

Russell, B. 1998 On our knowledge of general principles. In *Problems of philosophy*, 2nd edn, pp. 36–44. Oxford University Press. (Originally published in 1912.)

Sartwell, P. E. 1960 On the methodology of investigations of etiologic factors in chronic diseases—further comments. *J. Chronic Dis.* **11**, 61–63.

Schlehofer, J. R. & Zur Hausen, H. 1982 Induction of mutations within the host cell genome by partially inactivated herpes simplex type-1. *Virology* **122**, 471–475.

Susser, M. 1973 *Causal thinking in the health sciences*, pp. 41–47. New York: Oxford University Press.

Susser, M. 1977 Judgement and causal inference: criteria in epidemiological studies. *Am. J. Epidemiol.* **105**, 1–15.

Susser, M. 1988 Falsification, verification and causal inference in epidemiology: reconsiderations in the light of Sir Karl Popper's philosophy. In *Causal inference* (ed. K. J. Rothman), pp. 33–57. Chestnut Hill, MA: ERI.

US Surgeon General Advisory Committee on Smoking and Health 1964 *Smoking and health*. Washington, DC: US Department of Health, Education and Welfare (DHEW publication no. (PHS) 1103), US Public Health Service.

Vonka, V. (and 12 others) 1984 Prospective study on the relationship between cervical neoplasia and herpes simplex type-2 virus. II. Herpes simplex type-2 antibody presence in sera taken at enrolment. *Int. J. Cancer* **33**, 61–66.

Vonka, V., Kaňka, J. & Roth, Z. 1987 Herpes simplex type 2 and cervical neoplasia. *Adv. Cancer Res.* **48**, 149–191.

Weed, D. L. 1988 Causal criteria and Popperian refutations. In *Causal inference* (ed. K. J. Rothman), pp. 15–32. Chestnut Hill, MA: ERI.

Worall, J. 1985 Scientific discovery and theory confirmation. In *Change and progress in modern science* (ed. J. C. Pitt), pp. 301–331. Dordrecht, Germany: D. Reidel Publishing Co.

Wynder, E. L., Higgins, I. T. & Harris, R. E. 1990 The wish bias. *J. Clin. Epidemiol.* **43**, 619–622.

Yerushalmy, J. & Palmer, C. E. 1959 On the methods of investigations of etiologic factors in chronic diseases. *J. Chronic Dis.* **18**, 27–40.

Zur Hausen, H. 1991 Papillomavirus host cell interactions in the pathogenesis of anogenital cancer. In *Origins of cancer: a comprehensive review* (ed. J. Brugge, T. Curran, E. Harlow & F. McCormick), pp. 685–688. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.