Variance and Dissent

THE SURGEON GENERAL'S "EPIDEMIOLOGIC CRITERIA FOR CAUSALITY." A CRITIQUE

P. R. J. BURCH
Department of Medical Physics, University of Leeds, The General Infirmary, Leeds LS1 3EX, U.K.

(Received 4 October 1982)

Abstract—The methodology of the 1982 Report of the Surgeon General is examined with special reference to smoking and lung cancer. Part II of the Report describes the five criteria for causality that have guided the judgment of committees since 1964. I show that not one of the criteria, plausibly interpreted, is satisfied by the epidemiologic evidence for lung cancer. A weakness underlying all the Reports is a prior failure to recognize all the logical possibilities inherent in an association between smoking and a disease. The five criteria and the subjective method of "judgment" are inappropriate to a scientific analysis; they should be replaced by the objective testing of hypotheses. Limitations in the evidence and in concepts about tobacco carcinogenesis preclude definitive conclusions. Nevertheless, the entire association between cigarette smoking and lung cancer—at least in male Caucasoid populations—is unlikely to be explained by causation.

"The search for causes is perhaps the essence of science. The history of medicine reveals the many different causes that have been propounded to explain disease."


1. INTRODUCTION

In a Foreword, the Assistant Secretary for Health claims that the 1982 Report of the Surgeon General on "The Health Consequences of Smoking" presents... a comprehensive evaluation of the relationship between cigarette smoking and cancer." He concludes (Foreword, page v): "Cigarette smoking is the major single cause of cancer mortality in the United States. Tobacco's contribution to all cancer deaths is estimated to be 30 per cent."

The first report [2] by the Surgeon General in the series on smoking and health appeared in 1964 and it was followed in 1965 by Brownlee's critical review in the Journal of the American Statistical Association [3]. Brownlee argued that the Surgeon General's Committee had not established the case for causality between smoking and lung cancer. Neither did he believe that the causal hypothesis had been falsified; he held that it was not possible, at that stage, to reach definitive conclusions because, amongst other things, the genetic hypothesis had not been disproved.

No reference is made to Brownlee's analysis in the 1982 Report and, so far as I am aware, no extensive review of the Surgeon General's Reports and their methodology has been published since 1965. This paper, in common with Brownlee's, seeks to examine the fundamental assumptions and methods underlying the Reports with particular reference to the interpretation of the association between smoking and lung cancer. Alternative methods are proposed.
2. DISCUSSION OF "EPIDEMIOLOGIC CRITERIA FOR CAUSALITY"

After a brief introduction, Part II of the Surgeon General's 1982 Report, "Biomedical Evidence for Determining Causality", proceeds to "Epidemiologic Criteria for Causality". In the preamble to that section we soon reach the conclusion (p. 16): "Once an artifactual association has been ruled out, it is then necessary to determine whether the association is an indirect or direct (causal) one."

This key step at the early stage of the argument is incomplete and incorrect. Our choice is not of the either/or kind; "indirect" and "direct" associations are not mutually exclusive. Furthermore, another possibility has been overlooked. Given the existence of a genuine association between the habit (H) of smoking and the subsequent incidence of, or death from, a particular cancer (C) we are obliged by the rules of scientific inference to consider all the following possibilities:

(I) H causes C. (Some measure of caution needs to be exercised even at this juncture. The smoking habit might be associated with some other factor that provides the causal agent. For example, the means of ignition—matches, lighters—and not the combustion of tobacco might, in principle, be the source of the effective carcinogens. Other hypotheses of this kind will come readily to mind.)

(II) C, or an associated pre-C condition, causes H. (I have called this the "converse causal" hypothesis.)

(III) Some other factor causes, or predisposes to, both H and C. This is the "common cause" hypothesis; when discussed in genetic terms, it is often known as the "constitutional" hypothesis. On this view one or more genetic factors predispose to H; one or more genetic factors predispose to C; the association between H and C then arises at the genetic level.

(IV) Because (I) to (III) are not mutually exclusive, any combination of them might be necessary to account for an observed association.

Hypothesis IV seldom receives explicit recognition but its implications are profound and somewhat daunting. Thus, if any "evaluation" of an association is to be truly "comprehensive", the relative contributions of I, II and III, together with their respective confidence limits, would have to be assessed. No such assessments have appeared in the Surgeon General's Reports nor, so far as I am aware, in any others. My own attempts [4-6] to derive the magnitude of the causal component (I) of the association between smoking and lung cancer have been unsuccessful; errors of diagnosis and death-certification alone are apt to defeat such efforts.

Another implication of IV is that a weak, neutral, or even negative association between H and C does not rule out a causal connexion; a strong negative association between the genotypes predisposing to H and to C, for example, could overwhelm a weaker causal connexion. Needless to say, a positive association between H and C, however strong, is no proof of cause; we are required to dispose of II and III before opting for I. As we have seen, the failure of the Surgeon General's earlier Committee [2] to disprove hypothesis III was the main reason for Brownlee's rejection [3] of their conclusions.

The ethical and moral difficulties of randomizing for a "personal choice behaviour" such as smoking are stressed in the Report [1] and to ensure that an association..."is not due to a confounding variable, an entire body of data must exist to satisfy specific criteria, none of which by itself is an all sufficient basis for judgment" (p. 16). However, Rose and his colleagues have demonstrated the practicality of a randomized intervention trial in which one group of smokers, randomly chosen, is subjected to intensive advice to quit smoking while the other group, also randomized is not [7, 8]. Curiously, the Report fails to mention this exceedingly important investigation, which is one of the very few that was designed to discriminate between I and III. I refer to it in greater detail below.

The definition of "cause" (pp. 16-17) follows that of the 1964 report: it conveys..."the notion of a significant, effectual relationship between an agent and an associated
disorder or disease in the host". Later (p. 17): "the causal significance of an association is a matter of judgment which goes beyond any statement of statistical probability."

This emphasis on judgment, here and in other parts of the Report, is disturbing; subjective methods in science should be allowed as little scope as possible. If the relative contributions of I to III with their respective confidence limits are not made explicit, what reliance can be placed on (p. 63): “It is estimated that 85 per cent of lung cancer mortality could have been avoided if individuals never took up smoking”? In the absence of a quantitative demonstration, conforming to the accepted canons of scientific logic, we have little option but to treat such claims with scepticism; if we cannot place error limits on quantitative estimates we are probably not justified in making them.

The Report goes on (p. 17) to discuss the criteria for evaluating causal significance:

- a. The consistency of the association
- b. The strength of the association
- c. The specificity of the association
- d. The temporal relationship of the association and
- e. The coherence of the association.

These criteria, which form the basis of the Surgeon General's methodology, are unaltered from those adopted in the first Report [2]. They and their application to the association between smoking and lung cancer will be discussed in turn.

**Consistency of the association**

According to the Report: “This criterion implies that diverse methods of approach in the study of an association will provide similar conclusions.” And later (p. 17),...”replication assures that the association is not likely to be an artifact due to bias in the study methodology or subject selection, and that it is not indirect due to confounding variables such as diet, occupation, or genetics.”

How “similar” do conclusions need to be to conform to the consistency criterion? The Report offers us no guidance and in the sub-section on lung cancer (p. 21–62) the criterion loses all meaning. Thus Table 4 (p. 35) lists “relative risk ratios” (smokers vs non-smokers) for lung cancer mortality in 35 retrospective studies published over the period 1939–1970. For males, the observed ratios range from 1.2 to 36.0 and for females, from 0.2 to 5.3. Table 5 (p. 36) gives mortality ratios (cigarette smokers vs non-smokers) found in 8 prospective studies; these range from 3.76 to 14.2 in males and from 2.03 to 5.0 in females.

With ratios showing a range, overall, of more than two orders of magnitude it is not self-evident that any acceptable criterion of consistency has been satisfied. The Committee would appear to have been faced with the choice of either abandoning the causal interpretation, or of explaining away this enormous diversity. Instead, we are told (p. 34): “Regardless of the method, these studies have consistently found an association between smoking and lung cancer.” The fact that Table 4 gives two examples of relative risk ratios less than unity (they are negative associations) has either been overlooked, or negative and positive associations are both subsumed under “associations”. Might I suggest that ratios with, say, 95% confidence limits based on the numbers of deaths and age-standardization would assist evaluation and, when particular findings differ by some arbitrary degree from the overall weighted mean, comment would be appropriate? On a simple reading of Tables 4 and 5 any plausible criterion of consistency would appear to be overwhelmingly violated.

In connexion with the Hawaiian study [9] described by Hinds et al. in 1980 the Report notes (p. 34): “A significant positive association was found with cigarette consumption and lung cancer for all ethnic groups.” The association was indeed positive but not particularly impressive. In a simple regression analysis of age-adjusted rates, \( r^2 \), the simple coefficient of determination, was only 0.43 (it was 0.71 for beer consumption and lung cancer); in a multiple regression analysis for 10 ethnic-sex groups with the exposure variables: cigarettes, beer, wine and liquor, \( r^2 \) (the partial coefficient of determination, controlling for sex) was a mere 0.25 for cigarettes [9]. The findings for women of Hawaiian, Japanese and Chinese origin, with relative risks of 10.5, 4.9 and 1.8 respectively, were not published.
until 1981 and were evidently too late to be included in the Report, but Hinds et al. concluded [10]: “Cigarette smoking is clearly not the only cause, nor even the major cause of lung cancer in all populations of women.”

It is not without interest that Oriental populations generally show a much weaker association between smoking and lung cancer than Western populations (see below, under “Strength of Association”). There might be a good explanation for this contrast along causal lines, but if there is, readers of the Report should be made aware of it.

Data from the Swedish Twin Registry (not presented) are said to show “a significant excess of lung cancer” in the smoking twin of monozygotic pairs discordant for smoking (what is meant by “significant”? and the conclusion of the authors is quoted (p. 35): “The well-documented evidence of a causal association between smoking and lung cancer found in other subjects has been further supported.” Many authors have reservations about twin studies and, in common with most biological investigations, they are not free from complications [11]. Findings from the Swedish study [12] are illustrated here in Table 1.

### Table 1. Mortality from Lung Cancer by Cohort, Zygosity and Smoking-Discordant Categories, Sexes Combined

<table>
<thead>
<tr>
<th>Cohort Born</th>
<th>Dihydric pairs</th>
<th>Monozygotic pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No. of pairs</td>
<td>“Pooled low”</td>
</tr>
<tr>
<td>1886-1900</td>
<td>NA</td>
<td>0</td>
</tr>
<tr>
<td>1901-1925</td>
<td>1487</td>
<td>2</td>
</tr>
</tbody>
</table>

*Includes non-smokers and lower level smokers of smoking-discordant pairs.

where, it will be seen, small numbers alone impose severe interpretational restrictions. Data for the 1886-1900 cohort were regarded [12] as being too unreliable to be included in the main body of the report, but they were quoted, albeit briefly, in the notoriously unreliable context of lung cancer, and only in that context. The supposedly more reliable data for the 1901–25 cohort happen to be in agreement with the simplest expectations of the constitutional hypothesis but much larger numbers are needed for convincing statistical analysis. Large numbers, however, will not dispose of biological uncertainties. For example, do post-zygotic random changes sometimes occur to predispose one member of a monozygotic pair both to smoking and to lung cancer leaving the co-twin non-predisposed? It is a methodologic misfortune that studies of twins discordant for smoking can corroborate but cannot disprove the constitutional hypothesis.

**Strength of the association**

“Strength” and “consistency” are so interrelated it would be preferable to discuss them under the same heading, but the Surgeon General’s practice is followed here. We read, with a measure of incredulity (p. 17): “The relative risk ratio yields evidence on the size of the effect of a factor on disease occurrence and which, even in the presence of another associated factor without causal effect but coincident with the causal agent, will not be obscured by the presence of the non-causal agent.” This is followed by (p. 17): “A relative risk ratio measures the strength of an association and provides an evaluation of the importance of that factor in the production of a disease.”

These statements take for granted the size of the effect of a factor and the importance of that factor in the production of a disease. That is to say they assume in advance the causal hypothesis—hypothesis I above—and ignore II to IV. Strictly speaking, the strength of an association, in isolation, tells us nothing whatsoever about the “effect” of a “factor in the production of a disease”. A positive association, weak or strong, does not establish the causal hypothesis; and by the same token, a negative association does not reject it.

Nevertheless, the strength of the association, as determined in different populations, does provide us with valuable evidence that bears on our main quest. Suppose, for example, that studies in one population, P1, consistently yielded a relative risk, R1, and that studies in another population P2, with similar smoking habits, consistently yielded
a relative risk, R2. Then, if R1 and R2 differed significantly we should be justified in suspecting first, that the association is not wholly causal in origin and second, that the converse causal hypothesis (just possibly) and/or the common cause hypothesis (quite plausibly), should be invoked to account for the discrepancy between R1 and R2.

If, furthermore, we found that genetically-similar populations, say male Caucasoid, usually gave relative risk ratios of around 10 and that male Mongoloid populations gave relative risk ratios of between 1 and 4, we should be justified in suspecting that genetic differences might be responsible for these contrasting ratios, at least in part. In principle, such factors might act through the converse causal and/or the common cause pathways and we should require further evidence to discriminate between those two possibilities. Evidence relating to the effective invariance of the average age at onset of lung cancer in relation to different levels of smoking and the age at commencing to smoke in two male Caucasoid populations [13, 14] is inconsistent with the converse causal hypothesis [4, 5].

A situation approaching the foregoing is actually encountered: the Orient and the Occident generally show marked differences in relative risk. The findings [10] in Hawaii for women of Hawaiian, Japanese and Chinese origin have been cited above but data for men are equally suggestive. Table 5 of the Report (p. 36) shows that mortality ratios in seven predominantly Caucasoid populations (in Britain, N. America and Sweden) cluster around an average of about 10, with extremes of 7.0 (Sweden, based on a total of only 55 deaths) and 14.2 (Canadian veterans) but the value for Japan, based on 940 deaths, is given as 3.76. (Three significant figures seem rather generous.) Other findings in Oriental male populations (not cited in the Report) include the following: Chinese resident in Singapore gave a ratio of 3.8 [15]; in Northern Thailand a relative risk of 1.8 was obtained in a univariate analysis but this was reduced to 1.65 in a multivariate analysis that allowed for several "risk factors" and this value did not differ significantly from unity [16]; for mainland China a risk ratio of only 1.57 has been reported indirectly [17]. This last finding accounts, presumably, for the scepticism of Chinese scientists regarding the hazards of cigarette smoking claimed by most Western investigators. But again, I am bound to point out that if a strong correlation does not establish causation neither does a weak one reject it.

Returning to findings in women, the incidence of lung cancer in the Chinese in Hong Kong is among the highest in the world but the relative risk for smokers of manufactured cigarettes, vs nonsmokers, was found to be only 1.74; 53% of patients had never smoked cigarettes [18, 19].

It would be interesting to know whether these rather consistent ethnic-related differences can be explained by differences in smoking habits (including duration) or whether, as seems distinctly possible in view of the heavy smoking in Japan, biological differences might be responsible, at least in part.

The commonly observed difference between the sexes also calls for careful investigation. At the same level of smoking, mortality ratios (based on adequate numbers) are consistently lower in women than in men—see Table 6 of the Report (p. 38). Can this be explained in terms of the manner of smoking, or of the age of onset of smoking; or again, do we have to invoke biological differences of susceptibility? The carefully considered views of the Surgeon General's expert advisers would be welcome on these matters.

In discussing the high relative risk of lung cancer in smokers—in Caucasian populations, Oriental populations being ignored—the Report points out (p. 18): "To account for such high relative risks in terms of an indirect association would require that an unknown causal factor be present at least 10 times more frequently in the smokers and 20 to 30 more times more frequently among heavy smokers than among nonsmokers. Such a confounding factor should be easily detectable, and if it cannot be detected or reasonably inferred, the finding of such a strong association makes a conclusion concerning causality more probable." In spite of the great strides achieved in human genetics we still lack genetic markers—other than the disease itself—for many indubitably genetically-based disorders, including chronic diseases such as Huntington's chorea with a simple, autosomal heterozygous, form of genetic predisposition. It is unlikely that predisposition to any of the
several histological types of lung cancer is as simple as for Huntington's chorea (which is not entirely free from complications) and, so far as I am aware, there is no inherent reason why the association between smoking and lung cancer, in spite of its strength in Caucasoid populations, should not reside in the genome. I have given reasons for concluding that the association between smoking and ischaemic heart disease, with its interesting but readily interpreted marked age-dependence, is largely or wholly genetic in origin [20].

The section ends with: "Important to the strength, as well as to the coherence of the association, is the presence of a dose–response phenomenon in which a positive gradient between degree of exposure to the agent and incidence or mortality rates of the disease can be demonstrated." Hypothesis III also predicts a pseudo "dose–response" relation. To take the simplest postulates, smokers can be divided into two categories, social and habituated. Social smokers tend to be light smokers and to quit readily; habituated, genetically-predisposed smokers, tend to be heavy smokers. Hence, in any group of light smokers, social smokers will predominate and the association with lung cancer will be relatively weak; in any group of heavy smokers, habituated, genetically-predisposed smokers will predominate and the association with lung cancer will be strong. An apparent "dose–response" relation will be observed.

A "pure" causal hypothesis—in which the entire association is explained by hypothesis I free from biological interference—predicts the same response, from the same "dose", in different populations. A "pure" constitutional hypothesis predicts that the association between smoking and lung cancer depends, not on smoking levels, but on the strength of the associations between predisposing genotypes. Because on the constitutional hypothesis smoking levels in a population are likely to be determined to some degree by the frequency of smoking genotypes then some correlation—though not necessarily a very close one—between national mortality and national smoking levels would be expected.

From different but reasonable assumptions about the time relation between smoking rates, and deaths from lung cancer, the correlation between national mortality from lung cancer and national cigarette consumption is found to be weak [1, 4, 21]. To take an example of inconsistency, age-standardized mortality from lung cancer in Finnish men in 1960–61 was about double that in U.S. White men, whereas cigarette consumption in 1950 in Finland was about half that in the U.S. [4]. Many other anomalies of this kind exist—see Figs 9 and 10 (p. 43, 44) in the Report—and the pure causal hypothesis might, by this test alone, appear to be untenable. The existence of a weak correlation between national rates of mortality and smoking is consistent with a causal component but it is also consistent with a pure constitutional hypothesis and no causal action. However, for this type of evidence to be definitive it would be necessary to ensure comparability of death certification for lung cancer and, when testing the causal hypothesis, to have a good theory of the mechanism whereby cigarette smoking causes lung cancer so that explicit assumptions about temporal relations between the two could be adopted. One of the most unfortunate features of hypothesis I is that no good theory of mechanisms is available that is consistent with the salient features of epidemiologic evidence [5, 6].

According to the causal hypothesis the form of the "dose–response" relation—linear, quadratic, linear-quadratic, etc.—would be expected to be invariant in different populations. Many of us will have been impressed by the striking linearity of the graph of annual death rate vs average smoking rate in Doll and Hill's study of British doctors [22]. However, theory suggested that a multi-hit (at least 2-hit) mechanism of tobacco-carcinogenesis is necessary to account for the age-dependent features of lung cancer [23]. Doll and Peto [24] were able to show that a quadratic relationship gives a better interpretation of suitably analysed data for British doctors (based on 539 deaths) provided that the point for men smoking more than 40 cigarettes a day is ignored. Table 6 of the Surgeon General's Report (p. 38) shows a sub-linear "dose–response" relationship for Japanese males of all ages. Invariance of the form of the "dose–response" relation (or better, association) appears to be absent.

For those familiar with the literature of this subject Table 8 in the Report (p. 39), giving
mortality ratios by degree of inhalation in prospective studies, will occasion some surprise. Inhalation correlates with the daily level of smoking [25] and hence, in attempting to assess the effects of inhalation as such, it is essential to control for the level of consumption. The data in Table 8 for the ACS 25 State Study and the Swedish Study appear not to incorporate this elementary control and they are, therefore, highly misleading. In 1959 Fisher analysed [25] the retrospective data of Hill and Doll by daily rate of smoking (1-4, 5-14, 15-24, 25-49 and >49 cigarettes per day) and inhalation status (inhaler vs non-inhaler). Except for the group smoking 5-14 cigarettes per day, Fisher found that, within a given range of smoking, inhalers have a paradoxically lower risk of lung cancer than non-inhalers. He commented: "No particular importance need be attached to the test of significance. It disproves at about the 1 per cent level the hypothesis that inhalers and non-inhalers have the same cancer incidence. Even equality would be a fair knock-out for the theory that smoke in the lung causes cancer" [25]. In the 20-yr follow-up of British male doctors Doll and Peto standardized for age and amount smoked (in nine groups) and found, overall, that the risk of lung cancer in inhalers was 84% of that in non-inhalers [26]. These are remarkable findings and not easily accommodated to most causal theories; they surely warrant an extended discussion in the Report, especially in the light of Fisher's heavy sarcasm [25] on this issue.

Specificity of the association

This concept, which seems to be a hangover from Koch's (or Henle-Koch's) postulates, probably has little relevance to our present context. As the Report states (p. 18): "Specificity implies that a causal agent invariably leads to a single specific disease, an event rarely observed." In criticizing the 1964 Report [2], Brownlee pointed out [3]: "The Report accepts "specificity" as one of the criteria of "epidemiologic method", and in my opinion the way it claims the facts are in conformity with the criterion is to flatly ignore the facts." The section on specificity in the 1982 Report marks an advance (p. 19): "In summary, despite the fact that the demonstration of specificity in an association makes a causal hypothesis more acceptable, lack of specificity does not negate such an hypothesis, since many biologic and epidemiologic aspects of the association must be considered." If lack of specificity does not negate a causal hypothesis would it not be altogether more appropriate simply to delete "The specificity of the association" from the list of criteria for causal significance? Indeed, the lack of specificity revealed by the many positive and negative associations observed between smoking and various diseases have important implications that the Report ignores.

That cigarette smoking associates positively in some populations with malignancies of the bladder, prostate, pancreas, stomach, pharynx, larynx, esophagus, buccal mucosa and kidney does not, in my view, detract from the hypothesis that it causes lung cancer; indeed, those associations are consistent with the hypothesis that one or more components of cigarette smoke act as fairly general carcinogens with a wide distribution through the body. Negative associations between smoking and cancers are, however, another matter. Colorectal cancers tend to be negatively associated with cigarette smoking [27-29] and Choi et al. in a case-control study [30] found a marked negative association (p = 0.017) for central nervous system neoplasms (glioma, astrocytoma, glioblastoma and meningioma). These observations detract from the hypothesis that cigarette smoke acts as a general systemic carcinogen.

Particularly interesting is the consistently observed negative association between smoking and a non-neoplastic disorder, Parkinson's disease [26-28, 31, 32]. Nefzger et al. [31] and Kessler [32] found that diseases that are positively associated with smoking are infrequently reported in patients with Parkinson's disease, a finding that is readily accommodated to a genetic hypothesis. That smoking actually prevents Parkinson's disease seems unlikely. Westlund [33] pointed out that, on a prophylactic hypothesis, a secular increase in smoking would lead to a decrease in morbidity and mortality; and also, mortality from Parkinson's disease in rural areas would exceed that in urban areas. Norwegian statistics failed to confirm either expectation. Furthermore, differences in

Source: http://industrydocuments.library.ucsf.edu/tobacco/docs/rkkn0141
smoking habits are seen prior to the onset of Parkinson's disease and are therefore not caused by the disease itself [33]. There appears to be no realistic alternative to the hypothesis that the negative association between smoking and Parkinson's disease has a genetic basis [4]. If that should prove to be the case it will be surprising if at least some other positive and negative associations do not have a similar interpretation. A marked negative association ($p < 0.001$) has recently been reported between smoking and another non-neoplastic disease, ulcerative colitis [34]. Whether this results from preventive action (smoking on disease; disease on smoking), or from genetic factors, or both, remains to be determined.

**Temporal relationship of the association**

The Report is rather circumspect over this potentially important source of evidence; in the section on lung cancer only 32 lines of text (pp. 39-42) are devoted to it. By contrast, The Royal College of Physicians in 1971 went so far as to assert: “The chief reason for rejecting the genetic hypothesis is its inability to account for the enormous rise in death rates from lung cancer in the past half century” [35].

Under Epidemiologic Criteria for Causality (p. 19) we read: “The criterion of temporal relationship requires that exposure to the suspect etiologic factor precede the disease. Temporality is more difficult to establish for diseases with long latency periods, such as cancer.” Under Lung Cancer (p. 40): “One study examined the relationship between per capita tobacco consumption in 1930 and male lung cancer death rates in 1950 in 11 different countries.” The Report claims (p. 40) “... there was a strong positive correlation between tobacco consumption in 1930 and lung cancer death rates in 1950.” Figure 9 (p. 43) shows a graph of the data (why crude death rates?) with $r = 0.73 \pm 0.30$—a rather poor correlation, more readily connected with hypotheses III or IV, than with I. The data for U.S.A. and Great Britain serve to emphasize the weakness of the correlation: reading off from Fig. 9, a per capita cigarette consumption of about 1283 cigarettes per year in the U.S.A. in 1930 links with a crude male death rate of 194' deaths per million in 1950 but the corresponding statistics for Great Britain were 1139 cigarettes per year and 457 deaths per million. Large anomalies of this kind call for explanation, or, if none is forthcoming, an admission that the simple causal hypothesis cannot be sustained.

In any case such data do not provide an adequate test of a temporal relationship. Why not choose evidence in which time is presented as the independent variable with: (i) cigarette consumption; and (ii) lung cancer mortality, as dependent variables? Even from a crude analysis along these lines some interesting conclusions can be drawn [4, 5].

In the United Kingdom the sharp rise in cigarette consumption by women, which occurred after the First World War, lagged behind that of men by about 30 years [36, 37] (Fig. 1). However, when the rise in recorded mortality from lung cancer is studied in detail it is seen that the temporal pattern of increments, from one 5-yr period to the next, is remarkably synchronous in the two sexes from the beginning of the century to 1955 and then from 1965 onwards (Fig. 2 and [4, 5]). It follows that the main causes of the recorded increases in both sexes were also synchronous in both sexes and therefore could not have been cigarette smoking. There is little doubt that one important synchronous factor was improved diagnosis and better recognition of the disease, but other factors—oncogenic viruses, etc.—can by no means be ruled out [4, 5]. Unfortunately, the recorded increases in mortality in both sexes have been many times greater than would be expected from the hypothesis that the association between all types of smoking and lung cancer in the United Kingdom is wholly causal in origin [4, 5]. Nevertheless, if improvements in diagnosis were the same for both sexes the sex ratio of changes in mortality would compensate for in-phase diagnostic changes and reveal the temporal pattern of out-of-phase changes, including any caused by smoking. Accordingly, the temporal trends in the sex ratio of recorded mortality were compared with those calculated from changes in the consumption of tobacco of all kinds and a simple causal hypothesis [5]. The conspicuous lack of agreement between observed and expected trends (Figs 3 and D5 in Ref. [5]) might appear to reject the causal hypothesis decisively. However, incorrect assumptions about causal
relations, and/or a sex-differential in errors of death certification and tobacco consumption, might mask a causal (tobacco) component of change. Efforts to test hypotheses rigorously can readily be frustrated by defective data; reliable solutions to the enigma of cause remain elusive.

In another approach to temporal relations in England and Wales, Todd et al. [37] analysed the relation for males between: (i) cohorts that gave the highest lung cancer death rates over the period 1950 to 1967 in the six age groups 30-34 yr to 55-59 yr, and (ii) the cohorts that showed cumulative "constant tar" cigarette consumption equal to, or greater than, that of the cohort with the highest lung cancer rate in each age group. For one-hit initiating causal mechanisms we would expect the cohort with the highest cigarette consumption to be the one with the highest death rates. For all age groups that could be studied the cohort with the highest cumulative cigarette consumption was later, by 5-10 years, than the cohort with the highest mortality. Todd et al. doubted whether all these discrepancies could be plausibly explained by errors in the data [37].

Doll [38] proposed a model for the age-dependence of lung cancer in smokers and non-smokers which, on the basis of several tests, is readily shown to be untenable [5, 23]. My "precipitator" hypothesis of the causal action of cigarette smoke was devised to be consistent with the epidemiologic evidence that rejects Doll’s hypothesis; it has been tested using post-1950 temporal trends in sex- and age-specific cigarette consumption and lung cancer death rates for England and Wales [6, 39]. The hypothesis states that the risk of lung cancer is linearly related to the average rate of smoking r years before death. Three values of r, 2.5, 5 and 10 yr, have been tried. Only one out of 30 tests gave a significantly positive (but weak) association between the temporal trends in sex- and age-specific mortality and cigarette smoking [6, 39]. Although 29 out of 30 tests failed to support even a weak causal hypothesis the larger anomalies, especially in pre-1967 data, can probably be attributed, at least in part, to errors of death-certification.

Problems connected with errors of death-certification can be largely overcome by considering the temporal trends in mortality from all causes. When this is done, no support is given to the hypothesis that the (substantial) association between smoking and overall mortality in England and Wales is largely causal in origin [40].
It has to be conceded that the analysis of temporal trends has proved to be a largely sterile exercise inasmuch as it has failed to support any of several versions of causal hypotheses where lung cancer, ischaemic heart disease and overall mortality are concerned. The apparent rejection of causal hypotheses can sometimes be attributed to errors in the data although for ischaemic heart disease and overall mortality this excuse has limited scope [20, 40]. A major difficulty about tests involving temporal relations is that specific assumptions (or a range of specific assumptions) about causal mechanisms have to be adopted to make quantitative comparisons. Ideally, theories of mechanisms should derive from considerations that are independent of the epidemiologic evidence. Ad hoc hypotheses devised to fit the latter would be unsatisfactory; my "precipitator" hypothesis corresponds to my theory of the pathogenesis of acute infectious diseases [4].

Coherence of the association

"In order to establish the coherence of a specific association, other possible explanations for the association must be systematically considered and excluded or taken into account."

This prescription on p. 20 of the Report is impeccable; the execution (pp. 42–59 for lung cancer) pays scant attention to "other possible explanations".

The "dose–response relationship" is said (p. 42) to provide "great coherence with the known facts of the disease". However, as we have seen above, the form of the relationship is not invariant, ranging from sub-linear to quadratic, and international comparisons appear to be incompatible with a "pure" causal hypothesis; neither they, nor other aspects of "dose–response" relations, invalidate hypothesis III.

The next section, "Sex differences" (p. 42) begins with: "Males have had higher lung cancer death rates than females. This observation has been interpreted by some as contradictory to the causal role of smoking in lung cancer." References to Fisher and MacDonald are cited in support of this allegation. The reference to Fisher is inexplicable because he actually wrote (in 1957): "When the sexes are compared it is found that lung cancer has been increasing more rapidly in men relatively to women. The absolute rate of increase is, of course, obscured by improved methods of diagnosis, and by the increased attention paid to the disease, but the relative proportionate changes in men and women should be free from these disturbances, and the change has gone decidedly against the men."
But it is notorious, and conspicuous in the memory of most of us that over the last fifty years the increase of smoking among women has been great, and that among men (even if positive) certainly small. The theory that increased smoking is "the cause" of the change in apparent incidence of lung cancer is not even tenable in face of this contrast [25]. Thus, Fisher was discussing, not absolute death rates from lung cancer, but "relative proportionate changes" in mortality in the two sexes. The validity of his remarks about the relative increases with time in smoking and lung cancer mortality can be readily confirmed from Figs 1 to 3.

Analysis of the trends in mortality rates in the two sexes in England and Wales casts doubt on the claims of the Report on p. 44: "... the rise in female lung cancer mortality rates observed in the late 1950s and early 1960s [in the U.S.] appears to be reproducing the phenomenon noted among males 20 to 30 years earlier." When allowance is made for the approximately 30-yr separation between the sexes of the steep rises in cigarette consumption, the time-pattern of the change in death-rates in England and Wales for males, 1906–1946, bears no resemblance to that for females over the period 1936–1976 (Fig. 3). When no allowance is made for the interval between the rise in smoking in the two sexes the time pattern of the changes in death rates in men and women show, as mentioned above, a remarkable synchrony (Fig. 2 and Refs [4, 5]). In England and Wales, at least, recorded changes in death rates have been caused predominantly by factors unconnected with smoking.

"Lung cancer mortality and cessation of smoking" is considered on p. 45 but all the studies cited relate to mortality from lung cancer in self-selected non-smokers. No mention is made of the well-recognized principle that no valid conclusions regarding cause can be drawn from such studies of self-selected populations. The constitutional hypothesis predicts that quitters will tend to be non-habitual social smokers; studies of quitters before they gave up smoking, relative to continuing smokers, show that, on the average, these are dissimilar populations, with many differing features [41].

A greatly superior study design is offered by the randomized controlled intervention trial of the effects of quitting smoking. Such a trial is being conducted by Rose et al. [7] and results for a 10-yr follow-up have recently been reported [8]. Middle-aged (40–59 yr) male smokers at high risk of cardiorespiratory disease were allocated randomly: (a) to the "intervention" group which was subjected to intensive advice to give up cigarette smoking; or (b) to the "normal care" group which received no special advice over and above that encountered in ordinary life. The intensive advice was successful in reducing average levels of smoking in the intervention group well below those in the normal care group: after 1 yr, the average level of cigarette consumption in the intervention group was one-quarter of that in the normal care group; over 10 years the average nett reported reduction was just over one-half [8]. In the normal care group (731 men), 25 cases of lung cancer (deaths and registrations) were reported; in the intervention group (714), there were 22 comparable cases—a non-significant difference. Data for all deaths are free from diagnostic error, give the largest numbers, and are the most reliable: 123 deaths (17.2%) in the intervention group compared to 128 (17.5%) giving a (negligible) proportionate change on intervention of −2%, but with 95% confidence limits of −22% to +23% [8].

Probably the most curious finding, certainly the most unexpected although it was first noted at 5 years, is that for the category "all cancers other than lung cancer" several of which show, of course, a large positive association with smoking, the total number of cases (deaths and registrations) was 41 in the intervention group but only 19 in the normal care group. Rose et al. give p = 0.003 for this finding but rightly point out the difficulty of assigning a valid p to an a posteriori hypothesis [8]. They argue, I believe plausibly, that this apparent disadvantage of quitting smoking, although severe, is more likely to have been due to chance than to intervention. Further randomized intervention studies will be needed to resolve this serious issue. For the other four categories of mortality: lung cancer, coronary heart disease, "other causes", and total mortality, the null hypothesis is not violated, even remotely.

Source: http://industrydocuments.library.ucsf.edu/tobacco/docs/rkkn0141
Undoubtedly, the main lesson to be learned from this randomized study is that very large numbers of initial smokers will need to be studied if the null hypothesis—that quitting smoking has no effect on specific or all causes mortality—is to be disproved. The results for total mortality are entirely in line with the analysis of temporal trends of sex- and age-specific mortality from all causes in the whole of England and Wales [40], which failed to detect any causal influence of cigarette smoking when consumption was rising and no prophylactic influence when consumption was falling. Nevertheless, the controlled intervention trial is subject to a potential flaw even when large numbers are available: quitting smoking might be accompanied by other changes (dietary, stress, etc.) that could make the intervention group different from the normal care group, in factors other than smoking. Ideally, the entire “life-style” of all persons in such a trial should be monitored. Where specific causes of death are concerned the phenomenon of “detection bias”—the greater tendency of the physician to diagnose smoking-related diseases (especially lung cancer) in smokers than in non-smokers [42]—might also lead to false conclusions in the absence of careful necropsy examination of the causes of death. In the study of Rose et al., diagnosis in 88% of lung cancer cases in the normal care group and in 78% of the intervention group was verified by biopsy or necropsy examination and hence the scope for detection bias was limited. It is most unfortunate that this well-designed study—one of the very few that might have distinguished between causal and/or constitutional interpretations—has not been large enough to resolve the problem.

“Differences in lung cancer mortality by site of residence (urban versus rural)” is the topic discussed in the Report on pp. 45–47. Again, there is no acknowledgment that valid inferences about cause cannot be drawn from epidemiologic comparisons between self-selected, non-randomized groups. Mortality differences between urban and rural dwellers might be attributed to environmental and/or constitutional differences and the investigator should aim to test these hypotheses. Some interesting anomalies in urban–rural studies [43, 44] which ran counter to an environmental hypothesis, have been discussed elsewhere [4, pp. 353–4).

“Lung cancer mortality and occupation” (pp. 47–48) does not bear directly on the present discussion although it should be remarked that most studies of occupational groups are subject to the usual difficulties connected with self-selection.

The section, “Correspondence of lung cancer mortality among different populations with different tobacco consumption” (pp. 48–50) makes no mention of international contrasts, which, as pointed out above, embarrass pure causal theories; brief reference is made to findings among Mormons and California Seventh Day Adventists. That membership of these sects involves self-selection, or descent from self-selected progenitors, is overlooked. Findings for such groups are nevertheless interesting and call for careful examination. If Mormons conform to the dictates of their religion and refrain from smoking then, according to the pure causal hypothesis, their incidence of lung cancer should be the same as that in comparable non-smokers in the general population. Lyon et al. [45] compared the incidence in Mormons with that in non-Mormons (smokers and non-smokers) in Utah over the period 1967–75. The age-adjusted incidence of lung cancer in male Mormons was 46% of that in male non-Mormons and for females it was 44%. Details of the smoking habits of non-Mormons were not given but for males in the U.K. in 1958 [5], mortality from lung cancer in never-smokers was about 11% of that in all males; in a representative sample of U.S. Whites, aged 35–84 during 1966–68, mortality in never-smoking males was 19% of the whole sample and for White females it was 43% [46]. Hence, on the basis of the causal hypothesis the incidence of lung cancer in male Mormons appears to be at least twice that expected for a population of never smokers. On the constitutional hypothesis, the Mormon population—involving selection—comprises a mixture of never-smoking and smoking genotypes with, among males at least, a relatively high proportion of the former. From an extended analysis Enstrom concluded [47] that the cancer mortality patterns in Mormons are “not clearly explained by their smoking habits”.

“Lung cancer mortality and age-specific smoking patterns” are discussed (pp. 50–55)
with reference to changes in the prevalence of cigarette smoking among successive birth
cohorts of men and women in the U.S. As described above, under "Temporal Relationship
of the Association", Todd et al. [37] carried out a more detailed analysis, calculating
cumulative consumption of cigarettes by cohort in England and Wales; they found
considerable disagreement with expectations. If rates of smoking, as well as prevalence
data, are available it would be interesting to see a Todd-type analysis carried out on the
U.S. statistics. We see a hint that a similar discrepancy might well be uncovered. Thus,
the male cohort showing the highest prevalence of cigarette smoking is that of 1911-20
(Fig. 12, p. 51) whereas that with the highest age-specific death rates at 40 and 45 yr is
the 1930 cohort (Fig. 13, p. 52). Rates of smoking should also be considered; they might
or might not eliminate the discrepancy.

The Report comments (p. 50): "What appears to be a decline in lung cancer mortality
with age in the oldest age groups (75 years and older) is an artifact resulting from the
combination of cohorts with differing cigarette smoking exposures and mortality experi-
ences." It should be pointed out that this "artifact" has been present in the age-patterns
of lung cancer in England and Wales, for both sexes, ever since data have been
available—from 1901-05 onwards [4]. The modal age for age-specific death rates in men
remained almost constant at 67-68 yr from 1901-05 until 1951-55; in women it remained
at about 68-69 yr until 1926-30 then ranged from 72 to 76 yr from 1931-35 to 1966-70
[4]. I have proposed that a genuine and reproducible mode in the age-patterns of a disease
generally reflects the exhaustion with increasing age of the subpopulation predisposed to
that disease: as more and more predisposed persons develop the disease with increasing
age the number of predisposed unaffected persons in the sub-population diminishes and
absolute age-specific rates of onset (or death) eventually decline. As is well known, the
age-patterns of many neoplastic and non-neoplastic disorders exhibit at least one mode,
and sometimes more [4].

The section "Lung cancer mortality and premalignant changes in bronchial epithelium"
(pp. 55-59), prompts the question: Is the correlation between premalignant change and
cigarette smoking causal or non-causal? The question is not easy to answer because the
loss of cilia, for example, might be caused by smoking whereas carcinoma in situ might not.

"Epidemiologic Criteria for Causality" ends (p. 20) with the reiteration "The causal
significance of an association is a matter of judgment which goes beyond any statement
of statistical probability"; in connexion with particular cancers we are told: "the nature
of the association was assessed by applying the judgment criteria noted above. If all
epidemiologic criteria were judged to be satisfied and pathological and experimental data
are supportive, the term "causal" is applied to the association. The designation "major
cause" is used when the relative risk for the cancer in cigarette smokers is high." On
p. 62 the Report pronounces judgment: "Cigarette smoking is the major cause of lung
cancer in the United States." On p. 63, smoking is held responsible for 85% of lung cancer
mortality.

3. CONCLUSIONS

All the Surgeon General's Reports have adopted the same five epidemiologic criteria for
causality—quoted in Section 2 above—"no one of which is an all-sufficient basis for
judgment" (p. 17). They are utilized "for evaluation of the reported associations between
cigarette smoking and cancers of various sites in humans." Unfortunately, the criteria are
not given adequate definition but unless they are so lax as to be meaningless we can only
conclude that, in the context of lung cancer: (a) reported associations are inconsistent; (b)
the reported strength of association ranges widely; (c) the association has no specificity;
(d) the temporal relationship shows many anomalies; and (e), because of (a) to (d), together
with evidence relating, for example, to inhalation and to contradictions between experi-
mental and epidemiologic findings [3-5], the association lacks coherence. Because not even
one criterion is indisputably satisfied, it follows that the Report, on its own terms, should

Source: http://industrydocuments.library.ucsf.edu/tobacco/docs/rkkn0141
have rejected the causal interpretation of the association between smoking and lung cancer. It claims, however, that 85% of deaths from lung cancer are due to smoking. How does it do this? In Brownlee's words [3], which bear repetition: "the way it claims the facts are in conformity with the criterion is to flatly ignore the facts". This comment was made in 1965 with reference to (c), the specificity criterion, but in 1982 the criticism remains applicable not only to that but to all five criteria.

Does it then follow that the relation between smoking and lung cancer is wholly non-causal? In my view that inference would be as unjustified as the Report's own conclusion; on present evidence it seems very unlikely that the whole of the association observed, for example, in male Caucasoid populations should be attributed to causal effects of smoking but we cannot as yet rule out the possibility that some part should be. The Surgeon General's criteria aim at a dichotomous situation... "to determine whether the association is an indirect or direct (causal) one." This, as pointed out in Section 2, is a false dichotomy.

Probably the main methodologic weakness of all the Reports is the failure to appreciate and to expound all hypotheses I to IV. This initial failure is then compounded by inevitable errors of omission: the different kinds of epidemiologic evidence have not been exploited systematically in an effort to test these several hypotheses objectively. But, given the best scientific will in the world, a major barrier to effective hypothesis testing is the absence of a good and sustainable theory of the mechanism of tobacco carcinogenesis.

With respect to "possible mechanisms", Brownlee [3] remarked in 1965: "The difficulty with the smoking hypothesis is that it has not really got to first base on even lung cancer, let alone the other 24 causes of death, even though the matter has been under intensive investigation for ten years or more." In 1978, Oldham [48] was able to persist with a similar accusation: "The consequence is that, 28 years later, we still do not know how cigarettes cause lung cancer nor even, if we are particularly rigorous in our use of scientific logic, whether they do." Over the period 1968–1970 I slowly realised that a unified theory of growth and age-dependent non-neoplastic disease could also be applied to "natural" neoplastic diseases [49] and I subsequently hoped that it might illuminate the fundamental mechanisms of tobacco carcinogenesis. Together with Oldham [48] I am now far from certain that smoking does cause lung cancer but, until I tested it [6, 39], I believed that the "precipitator" hypothesis of mechanism was the most promising. My failure to corroborate that hypothesis, even in a weak form, might be attributed entirely to errors in the data. Nevertheless, this prolonged inability of many investigators to formulate a quantitative theory of tobacco carcinogenesis that is both internally consistent and in conformity with the more dependable features of the epidemiologic evidence, would be less surprising if the association is largely or wholly constitutional in origin.

Turning aside from theoretical problems, inaccurate data constitute, perhaps, the second major barrier to rigorous hypothesis testing. Comparisons between death certification and post mortem necropsy studies in the U.S. and U.K. have revealed large discrepancies [40, 50–52]. Clearly, if we are ever to determine with confidence the relative contributions of I and III to an overall association in a given population we shall need, at the least, virtually error-free diagnosis and death-certification. That will not guarantee accurate conclusions but it is a necessary precondition for them. On the other side of the equation, we shall also need accurate statistics for cigarette consumption and, in the final analysis, we might even need to know the time-concentration relation for effective carcinogens at the cells or sites at risk.

Finally, if any author or committee is so bold as to proclaim that x% of cases of lung cancer, heart disease, etc. are caused by cigarette smoking we should insist on the error limits of x and an explanation as to how they were derived. If the derivation conforms to the recognized rules of scientific and statistical inference then the estimate may be accepted. But if no error limits can be satisfactorily quoted such claims deserve to be treated with scepticism even should they appear under the imprimatur of the Surgeon General.
REFERENCES