STATISTICAL CRITICISM

IRWIN D. J. BROSS, PH.D. *

IN THE GREAT debate over smoking and lung cancer the quality of statistical criticism was, I think, rather poor (despite the eminence of the critics). The bitter lesson to be learned was this: *The "rules of the game" for statistical criticism need to be spelled out more clearly and completely.* These remarks represent a first step in this direction. While I will draw my examples from the lung cancer debate, similar instances can be found in most of the scientific areas in which statistical methods are employed. Much that nowadays passes as statistical criticism is superficial and sophomoric in character and serves to obscure a scientific discussion rather than to clarify it.

Let me emphasize at the start that the purpose of ground rules is to put the statistical critic on his mettle—not to muzzle him. The rules of the game for the *proponent* of scientific hypothesis, discussed in various texts on statistics and logic, can help him to make the statements that are warranted by the data. If the proponent flagrantly violates these rules, they provide a basis for calling him to account.[4] In the same way, ground rules for a critic will help him to distinguish valid objections. Of course, they also provide a basis for calling a critic to account for irresponsible attacks on scientific study. If both proponents and critics have to watch their P's and Q's, we might hope that it would be easier to achieve broad agreement on scientific issues.

THE ROLE OF A CRITIC

As a first step toward the ground rules of statistical criticism, let us examine the roles of the critic and the proponent. In what follows, the critic will be considered as opposed to the proponent in the sense that he denies the proposed scientific hypothesis or at any rate denies that it has been demonstrated.

Although the critic's role appears purely negative, it has a positive side to it. Implicitly (and sometimes explicitly) he puts forth a counterhypothesis. This point may be clarified by a simple example. Let us say that a critic objects to the conclusions of a scientific study because the proponent has not used significance tests. This objection would be trivial if, for example, the value of the chi square was actually enormous. However, it would be a strong objection if a difference between two series (which was essential to the proponent's argument) was not significant when the test was performed. But why is this objection strong? Because the critic can now frame a tenable counterhypothesis that explains the results in terms of sampling variation alone. Since the proponent cannot rule out this counterhypothesis, he cannot establish his own hypothesis.

In much the same way a critic who objects to a bias in the design or a failure to control some established factor is, in fact, raising a counterhypothesis (even though he may not state it). Since the counterhypothesis is essential in the logical structure of criticism, it facilitates debate when it is explicitly stated.[2] When the hypothesis is so stated, the basic question suggests itself: What is the responsibility of a critic with respect to his counterhypothesis?

A CRITERION FOR CRITICISM

Consider the following tentative rule: The critic has the responsibility for showing that his counterhypothesis is tenable. In so doing, he operates under the *same* ground rules as a proponent.

This rule may appear to conflict with the principle that the *burden of proof* rests on a proponent, but this is not the case. Although both critic and proponent may operate under the same rules in establishing their respective hypotheses, there is a great difference in what happens next. When a critic has shown that his counterhypothesis is tenable, his job is done (while at this point the proponent's job is just beginning). A proponent's job is not finished as long as there is a tenable hypothesis that rivals the one he asserts.

Many critics seem to employ a rule that is

From the Cornell University Medical College, New York, N.Y.
* Present address: Department of Statistics, Roswell Park Memorial Institute, Buffalo, N.Y.
Received for publication May 15, 1959.

much weaker than the tentative rule stated. They feel that a critic's responsibility ends when he merely *presents* a counterhypothesis without showing it to be tenable. This I regard as unrealistic because it imposes an impossible task on a proponent. He would be required to rule out every *conceivable* hypothesis. Since there are an unlimited number of such hypotheses, there would be no end to the proponent's labors. By restricting consideration to *tenable* hypotheses, the proponent's task becomes feasible (although onerous).

Tentative rule does not impose any impossible task upon the critic, since he can employ the usual scientific procedures to show that his counterhypothesis is tenable. For example, a minimal requirement would be that the effects predicted from the critic's hypothesis should be in line with the actual data, at least in direction and order of magnitude. The additional arguments needed to establish tenability depend on the nature of the hypothesis. For instance, if the hypothesis involves sampling variation alone, it would be tenable in any study employing samples. For hypotheses involving an artefact, the experience with this artefact in previous studies can be used to establish the direction and magnitude of the effect. It may even be possible to show the effect operating in the proponent's data. When a counterhypothesis involves a well known real factor, e.g., age or sex in an epidemiological study, it would be sufficient to mention the relationship, e.g., death rates from cancer tend to increase with age. However, when a counterhypothesis is novel or controversial, a critic (like a proponent in the same circumstances) will have to develop a strong argument.

For these reasons, the suggested rule for criticism seems to be both fair and feasible (and I will employ it in criticizing the critics).

HIT-AND-RUN CRITICISM

The bulk of statistical criticism is of the hit-and-run variety—the critic points out some real or fancied flaw and supposes that his job is done. Indeed, some critics appear to labor under the misconception that if some flaw can be found in a study, this automatically invalidates the author's conclusions. Since the critic makes no attempt to develop a tenable counterhypothesis, his performance is on a par with that of a proponent who glances at his data and then jumps to his conclusion.

Two examples should suffice to make this point plain. Quite a number of the critics of the Hammond and Horn study[9] (along with other prospective studies) have called attention to the possibility of misclassification of the cause of death on death certificates. Most of these critics dropped the matter at that point (apparently under the impression that they had scored a hit). However, if they had followed the usual scientific procedures in developing a tenable hypothesis—if they had looked at other studies or existing theory[3]—they would have found that misclassification tends to *diminish* observed differentials. If they had compared the Hammond and Horn tabulations of "purified" data, i.e., cases with confirmed diagnosis, with the "unpurified" data, they would have *seen* how misclassification operated to reduce differentials.

Another example is Berkson's model for a selection bias based on an item in the protocol of the Hammond and Horn study (initially sick individuals were excluded).[1] The model itself is a good example of constructive criticism, since it formulated the objection in a precise fashion that facilitated both theoretical and empirical investigation. However, I was amazed when, in talking with several statisticians, I encountered the opinion that this *model* seriously jeopardized the Hammond and Horn conclusions. Of course a model carries no weight in a scientific argument until it has been shown to be tenable. This particular model was untenable because it predicted that the differentials found in the study would *shrink* as time went by, whereas, if anything, the change with time was the *opposite* direction. What is more, even if the model had been tenable it would have been of little value in a counterhypothesis, since it could be shown mathematically that the bias could produce only slight differentials (the observed differentials had a different *order of magnitude*).[4]

We see, then, that it is not enough to spot flaws in a study; a responsible critic would go on to show how these flaws lead to a counterhypothesis that can explain the observations. If a critic fails to build a tenable hypothesis, he clearly fails in his duty.

DOGMATIC CRITICISM

To show that his counterhypothesis is tenable, a critic may use arguments based on

current statistical principles and practices. However, a critic has no license to make exaggerated claims, unfounded assumptions, or dogmatic assertions (even if the statements are quoted from statistical textbooks).

Consider the following quotation from Sir R. A. Fisher, which has been echoed by other eminent critics: "The evidence linking cigarette smoking with lung cancer, standing by itself, is inconclusive, *as it is apparently impossible to carry out properly controlled experiments with human material.*"[10] (The italics are mine.)

This blanket condemnation rests largely on one defect of the prospective studies as compared to controlled experiments. The exposure to cigarette smoke, i.e., smoking habits, is determined by the personal choice of each individual, whereas ideally the exposure would be set by the experimenter (using a randomized allocation). Because of the lack of randomization, there is a *potential* "self-selection" bias (which suggests a counterhypothesis). If this counterhypothesis can be rendered tenable, then, indeed, the proponent's evidence is "inconclusive."

Instead of attempting to make the self-selection hypothesis tenable, Fisher simply dismissed the entire body of epidemiological data (involving carefully collected information on hundreds of thousands of individuals). He did so on the basis that the data do not meet certain theoretical standards for "properly controlled experimentation." This seems to me a gross violation of the empirical spirit of modern science *and* of modern statistics. It raises the theory of statistics, e.g., randomization, to the level of dogma.

SPECULATIVE CRITICISM

While I do not agree with those who say that there is no place for speculation in a scientific article, I do feel that there are definite restrictions on hypothetical excursions. For one thing, speculation should be clearly labeled as such; for another, speculations should not enter the conclusions. These restrictions apply equally to proponent and critic.

There is one type of counterhypothesis in which the temptation to speculate is very strong—hypotheses based on a new "real world" factor. A statistician should be especially careful with this type of substantive hypothesis because he is now in the domain of the subject matter field—he is functioning as an epidemiologist or sociologist or psychiatrist (depending on the nature of the new factor) rather than as a statistician.

The task of establishing the tenability of a substantive counterhypothesis is more difficult than that for a methodological counterhypothesis, since "local" ground rules, i.e., those of the particular scientific field, come into play. For example, in epidemiology a proposed new factor has to be consistent with the broad incidence patterns of the disease, e.g., geographic distribution, time trends, and sex ratios.

While numerous substantive counterhypotheses have been introduced in the lung cancer controversy, there has been practically no attempt to render such hypotheses tenable. Thus Berkson brought up Pearl's[12] "rate of living" hypothesis[2] but frankly admitted that: "Actually I do not know of any independent evidence for such an effect of smoking." He also cited the "constitution" hypothesis, noted one of its shortcomings, and remarked: "I do not profess to be able to track out the implications of the constitutional theory or to defend it. . . ." While it is to Berkson's credit that he clearly labeled these two counterhypotheses as speculative, they appear to play an important role in his subsequent rejection of the "carcinogen" hypothesis, i.e., speculations enter his conclusion.

It may be argued that it is too stringent to require a critic to show that his substantive counterhypothesis is tenable because he is not actually *asserting* it but merely *suggesting* it as a possible line for future research. However, I fail to see how a critic contributes to the scientific process if the suggested avenue for research is, in fact, a dead end road. Nor can I see how a critic can expect to point out a sensible direction for research unless he explores the tenability of his counterhypothesis —for example, by checking whether his notion jibes with the incidence pattern for lung cancer.

The most striking feature of lung cancer incidence is the drastic increase in the age specific male death rates over the past generation. This rapid increase is virtually *unique*—the female death rate shows much less change, other cancer rates are fairly stable, and the rates for other causes of death either show relatively minor changes or else are rapidly *decreasing*. The peculiar behavior of the male lung cancer rates poses some difficult ques-

tions for any substantive hypothesis. Why is lung cancer thus singled out? Why are male death rates affected and not female death rates? Why should this have happened in the last generation? I leave it to the reader to put these questions to some of the counterhypotheses raised, e.g., those based on "stress," "genetic factors," and "constitution."

In my opinion, even a cursory exploration would have shown most of the critics that their substantive counterhypotheses were untenable. Had this been done, much of the confusion in the lung cancer debate would have been avoided.

TUBULAR CRITICISM

Proponents of scientific hypotheses are often justly criticized for their "tubular vision"—a remarkable inability to "see" the evidence unfavorable to their hypothesis. Critics are equally subject to this type of defective vision. For example, Berkson[2] complained that "virtually all of the evidence" that cigarette smoke is carcinogenic comes from epidemiological-statistical studies. He was unable to "see" the evidence from vital statistics, combustion chemistry, animal experiments, lung tissue pathology, etc.

Tubular vision also occurs in the examination of actual data. Since Berkson is one of the few critics who (1) dealt with data, (2) stated his counterhypotheses, and (3) made a serious effort to establish their tenability, I will draw my examples from his work.[2] However, judging *over-all* performance, I would say Berkson far excels the other critics.

To appreciate the illustrations, we first must understand the *purpose* of Berkson's analysis of the Doll and Hill data.[5] His counterhypothesis was: "The observed associations are 'spurious' . . . the result of the interplay of various subtle and complicated 'biases.'" To establish tenability, Berkson first undertook to show that ". . . there can hardly be any doubt that association is shown for 'all or nearly all' causes of death. . . ." in prospective studies. Before examining Berkson's arguments for this crucial point, let us see how it is used to establish the counterhypothesis. Berkson said: "For myself, I find it quite incredible that smoking should cause all these diseases. . . . And if we are not crassly to violate the principle of Occam's razor, we should not attribute to each separate association a radically different explanation."

I would not interpret Berkson's remarks as a denial that an environmental factor, e.g., polluted milk, can be responsible for more than one disease. Hence tobacco smoke, which is chemically quite complex (containing nicotine, polycyclic hydrocarbons, etc.), *might* induce or aggravate several different diseases (e.g., lung cancer, coronary thrombosis, or chronic bronchitis) by radically different "specific" etiological mechanisms. We can, however, make a distinction between those diseases in which an etiological hypothesis based on chemical components in tobacco smoke can be supported by independent evidence (call these "specific" diseases) and the many other causes of death in which a corresponding hypothesis would be highly speculative (call these "nonspecific" diseases). Now *if* we find that many of the "nonspecific" diseases are associated with smoking, then I quite agree with Berkson that the "simple" hypothesis of a general bias running through the data is clearly preferable to the "complex" hypothesis requiring a large number of speculative hypotheses to account for the associations. Moreover, if we also find that the bias effect is similar in direction and magnitude to the effects found for the "specific" diseases, then we have a tenable counterhypothesis for the whole of the data and the proponents of "specific" hypotheses are in a hopeless position.

Of course, this argument hinges on a demonstration that there is "generalized association" in the "nonspecific" diseases. For this purpose, Berkson started with a Doll and Hill tabulation (Table 29)[2] that gave the death rates in 4 tobacco consumption categories for "lung cancer," "coronary thrombosis," "other respiratory diseases," "other cancers," and "other diseases."

For a significance test of "generalized association," Berkson suggested that: "Appropriate here is some form of permutation test. . . ." He went on to say that: "However it is figured, the probability of getting by chance . . . consistently higher death rates among the heavy smokers than among any of the 3 categories of less than heavy smokers, in each of 5 predesignated categories of cause of death, and in agreement with the independently obtained similar finding in the prospective study of Hammond and Horn,[9] must be considered negligible."

Here, I think, is an instance of "tubular vision." Two of the 5 categories represent "spe-

TABLE 1

DEATH RATES FOR VARIOUS SMOKING CLASSES FOR INDIVIDUALIZED CATEGORIES OF DISEASE (REPORT OF DOLL AND HILL[5])*

Category	No. deaths	Death rate, standard./1,000 Non-smok.	Men smoking a daily average of: 1-14 gm.	15-24 gm.	25+ gm.
Cancer					
Lung	84	0.07	0.47	0.86	1.66
Up. respir. & digest. tract	13	0.00	0.13	0.09	0.21
Stomach	32	0.41	0.36	0.10	0.31
Colon & rectum	57	0.44	0.54	0.37	0.74
Prostate	30	0.55	0.26	0.22	0.34
Other sites	88	0.64	0.72	0.76	1.02
Respir. dis.					
Pulm. tuberculosis	19	0.00	0.16	0.18	0.29
Chron. bronchitis	42	0.12	0.29	0.39	0.72
Other respir. dis.	65	0.69	0.55	0.54	0.40
Coronary thrombosis	508	4.22	4.64	4.60	5.99
Other cardiovas. dis.	279	2.23	2.15	2.47	2.25
Cereb. hemorrhage	227	2.01	1.94	1.86	2.33
Peptic ulcer	18	0.00	0.14	0.16	0.22
Violence	77	0.42	0.82	0.45	0.90
Other dis.	183	1.45	1.81	1.47	1.57

*Data from Table 34 of Berkson.[2]

cific" diseases while a third, "other respiratory diseases," largely reflects the influence of chronic bronchitis—another "specific" disease. In other words, most of the evidence that Berkson used to *deny* "specific" effects came from these very effects! Indeed, *unless* these "specific" effects are included, there is little evidence for a "generalized association" in Table 29. Thus, while a permutation type test is significant at the 5% level for the 2 "specific" diseases, the corresponding test for the 2 "nonspecific" causes is definitely *not* significant.

TUBULAR CRITICISM AND DATA

Berkson himself did not seem satisfied with his inferences from Table 29 for he proceeded to construct (from Doll and Hill tabulations), Table 34, which listed 15 causes of death (and hence permitted segregation of "specific" causes). This table is reproduced as Table 1 in this article.

Berkson clearly "sees" his "generalized association" operating in Table 1, but the only analytic evidence offered is: "The death rate for heavy smokers is higher than that for nonsmokers in 12 of the 15 categories, although in several instances the number of deaths, the differences of rates, or both, are small."

A "permutation" test that could be used on this evidence is the sign test. [Strictly speaking, the death rates in the different causes may not be independent because overenumeration in one cause might lead to underenumeration in a related cause.] Let I be the total number of "inversions" (i.e., cases in which the death rate was lower for the heavy smokers than for the nonsmokers). Let NI be the total number of "noninversions." Then, using the sign test:

$$\frac{[|NI-I|-1]^2}{NI+I}=\frac{[|12-3|-1]^2}{12+3}=\frac{64}{15}=4.27$$

Since the 5% level critical value is 3.84, the sign test is significant, and we would reject the null hypothesis that sampling variation alone can account for the result cited by Berkson. Unfortunately, a departure in the observed direction might be due to either "specific" or "generalized" association, and if we try to subdivide the causes the numbers will be so small that the sign test will have little power.

However, we can call on the big brother of the sign test, the sequence sign test, to do the job. As before, we count "inversions" but this time we consider all 6 of the pairwise comparisons that can be made between the 4 death rates for each cause. In Table 1 we count a pairwise comparison as an "inversion" if the death rate to the right of the other member of the pair is the smaller one. Thus, for cerebral hemorrhage the sequence of rates is 2.01, 1.94, 1.86, and 2.33. Starting with 2.01 as the "left hand" member of the pair, we have "inversions" for 1.94 and 1.86 and a "noninversion" for 2.33. Moving on to 1.94 as the "left hand" member of the pair, we have an "inversion" for 1.86 and a "noninversion" for 2.33. Finally for the pair 1.86, 2.33 is a "noninversion." So for this cause we have 3 "inversions" and 3 "noninversions." Table 2 lists the results for the causes in Table 1.

For all causes, there are 27 inversions and 63 noninversions and the sequence sign test is:

$$K\frac{(|NI-I|-1)^2}{NI+I}=\frac{9}{13}\,\frac{(|63-27|-1)^2}{90}=9.42$$

where

$$K=\frac{9}{2(\text{no. factor categories})+5}=\frac{9}{2(4)+5}=\frac{9}{13}$$

and represents an adjustment for the fact that the 6 pairwise comparisons in a cause are not independent. This test is significant at the 1% level.

Since the sequence sign test is more powerful than the sign test, we are now able to segregate the "specific" and "nonspecific" causes (Table 2). I have also separated off 3 causes (subtotal B in Table 2) that are of questionable value for our purposes. There are less than 20 deaths in each of these series, and they were tabulated separately only because "specific" effects were suspected.

The "specific" causes (subtotal A in Table 2) show up as highly significant (8.65), although there are only 3 of them. The "nonspecific" causes in toto show 25 inversions and 29 noninversions, which is close to the expected values under the null hypothesis, i.e., 27 and 27, and is, of course, not significant (0.12). We might expect a "generalized association" to show most clearly in the "nonspecific" major causes since these 3 causes account for 40% of all the deaths, but this is not the case (subtotal D in Table 2). Again in view of Berkson's demand for an ". . . explanation for the association shown with cancers . . . of such sites as the colon, stomach, and pancreas . . ."[2] we might expect a striking result for "nonspecific" cancers, but this was not found (subtotal C in Table 2). In short, a permutation type analysis fails to detect Berkson's "generalized associations" in the Doll and Hill data (although it picks up associations for the "specific" diseases easily enough).

TABLE 2

INVERSION COUNT FOR TABLE 1 (BY CAUSE)*

Gp.	Invers.	Noninvers.
"Specific" causes		
Lung ca.	0	6
Chron. bronchitis	0	6
Coronary thrombosis	1	5
SUBTOTAL A	1	17
Questionable causes		
Pulm. tuberculosis	0	6
Up. respir. ca.	1	5
Peptic ulcer	0	6
SUBTOTAL B	1	17
SUMMARY TOTALS		
All "nonspecific"	25	29
All causes	27	63
"Nonspecific" ca.		
Stomach	5	1
Colon & rectum	2	4
Prostate	4	2
Other sites	0	6
SUBTOTAL C	11	13
"Nonspecific" major causes		
Cereb. hemorrhage	3	3
Other cardiovas. dis.	2	4
Other dis.	2	4
SUBTOTAL D	7	11
Other "nonspecific" causes		
Violence	1	5
Other respir. dis.	6	0
SUBTOTAL E	7	5

*Table 1 is given as Table 34 by Berkson.[3]

No permutation analysis is presented in Berkson's article[2] and instead he simply cited 2 examples of nonspecific causes that he believed show evidence of association. Berkson had reservation about 1 of these causes, "violence," but the other ("cancer: other sites") is "notable": ". . . this group shows a graded increase of death rate with increased amount of smoking, from a rate of 0.64 for nonsmokers, to a rate of 1.02 for heavy smokers." However, as can be seen from Table 2, the cited example is the *only* "nonspecific" cause whose rates show a graded increase (I equals 0 and NI equals 6) with increased smoking. If this one cause is to be regarded as a strong argument for generalized association, what are we to make of the category "other respiratory diseases" that shows a very similar pattern but in the *opposite* direction (I equals 6 and NI equals 0)?

SUMMING UP

The lengthy illustration of "tubular vision" in the examination of data contains several important lessons for statisticians. First, it shows how dangerous it is—even for an experienced and competent statistician—to draw inferences by scanning data and picking out favorable cases. Second, it indicates how analytic tools can help to safeguard against the "tubular vision" to which we are all liable. Third, Berkson's approach[2]—while not successful for the Doll and Hill data—illustrates how an argument for the tenability of a counterhypothesis can be developed from a proponent's own data. Fourth, the example shows that the task of a responsible critic can be as difficult and exacting as that of a responsible proponent (whereas "hit-and-run" criticism is child's play).

In my discussion of the role and responsibility of the statistical critic, my theme has been: we should not have a "double standard" in science and statistics, one standard for proponents and another for critics. The same

ground rules should apply to both. If a proponent should not jump to his conclusions or base them on dogma or speculation, neither should a critic. If a proponent should be wary of "tubular vision," so should a critic. In short, we might frame the following "golden rule" for critics: Do unto a proponent as you would have *him* do.

TECHNICAL APPENDIX

Derivation of the sequence sign test follows directly from a result given on page 241 of Feller's *An Introduction to Probability Theory and Its Applications*.[6] Feller proves that for a single sequence (under the null hypothesis) the number of inversions is asymptotically normally distributed with a mean

$$(E_1)=\frac{n(n-1)}{4} \quad \text{and variance}$$

$$(V_1)=\frac{(2n+5)(n)(n-1)}{72} \text{ or } \frac{(2n+5)(E_1)}{18}.$$

Assuming independence for M causes, we find at once that the total number of inversions (I) is asymptotically normally distributed with

$$E=M(E_1), \text{ and } V=M(V_1). \quad \text{Since } \frac{(I-E)^2}{V}$$

is asymptotically distributed as chi square with 1 degree of freedom, the sequence sign test (uncorrected) follows when the substitution $I+NI=2E$ is made. I have included a correction for continuity analogous to the one used in the sign test.[11]

REFERENCES

1. BERKSON, J.: Statistical study of association between smoking and lung cancer. *Proc. Staff Meet. Mayo Clin.* 30: 319-348, 1955.

2. BERKSON, J.: Smoking and lung cancer; some observations on 2 recent reports. *J. Am. Statist. A.* 53: 28-38, 1958.

3. BROSS, I. [J.]: Misclassification in 2 × 2 tables. *Biometrics* 10: 478-486, 1954.

4. BROWNLEE, K. A.: Note on effects of nonresponse on surveys. *J. Am. Statist. A.* 52: 29-32, 1957.

5. DOLL, R., and HILL, A. B.: Lung cancer and other causes of death in relation to smoking; second report on mortality of British doctors. *Brit. M. J.* 2: 1071-1081, 1956. Cited by BERKSON, J.[2]

6. FELLER, W.: An Introduction to Probability Theory and Its Applications, 2d ed., vol. 1. New York, N.Y. John Wiley & Sons, Inc. 1957.

7. FISHER, R. A.: Cancer and smoking. [Letter to the Editor.] *Nature, London* 182: 596, 1958.

8. FISHER, R. A.: Lung cancer and cigarettes? [Letter to the Editor.] *Nature, London* 182: 108, 1958.

9. HAMMOND, E. C., and HORN, D.: Relationship between human smoking habits and death rates; follow-up study of 187,766 men. *J. A. M. A.* 155: 1316-1328, 1954. Cited by BERKSON, J.[2]

10. L[AURENCE], W. L.: Cigarette-cancer links disputed. *New York Times* 107: (sect. 4) 7, Dec. 29, 1957.

11. MOSTELLER, F.: Some statistical problems in measuring subjective response to drugs. *Biometrics* 8: 220-226, 1952.

12. PEARL, R.: The Rate of Living; Being an Account of Some Experimental Studies on the Biology of Life Duration. New York, N.Y. Albert A. Knopf, Inc. 1928. Cited by BERKSON, J.[2]

13. WYNDER, E. L.; BROSS, I. J.; CORNFIELD, J., and O'DONNELL, W. E.: Lung cancer in women; study of environmental factors. *New England J. Med.* 255: 1111-1121, 1956.