

## TEST OF THE LINEAR–NO THRESHOLD THEORY: RATIONALE FOR PROCEDURES

**Bernard L. Cohen** □ University of Pittsburgh

□ A tightly reasoned justification is presented for the procedures used in our test of the linear-no threshold theory of radiation carcinogenesis by comparing lung cancer rates with average radon exposure in U.S. counties. A key point is its dependence on ecological variables rather than on characteristics of individuals and the principal problems involve treatment of potential confounding factors (CF). The method of stratification is introduced and shown to be preferable to multiple regression for evaluating effects of confounding. Utilizing numerous available CF reduces the problem of representing a complex confounding relationship by the average value of a single CF. The requirements on a CF for affecting the results are quantified in terms of its correlations with lung cancer rates and radon levels and it is shown that the existence of an unknown confounder satisfying these requirements is highly implausible. Effects of combinations of confounding factors are treated and shown not to be important. The problem of confounding factors on the level of individuals is resolved. Consideration of plausibility of correlations is used in several applications, including treatments of uncertainty in smoking prevalence, within county differences in radon exposure between smokers and non-smokers, variations in intensity of smoking, differences between measured radon levels and actual exposures, etc. Examples are presented for all applications. The differences between our study and case-control studies, and the advantages of each for testing the linear-no threshold theory, are discussed.

*Keywords:* linear-no threshold, radiation carcinogenesis, confounding, stratification, plausibility of correlation, dose-response

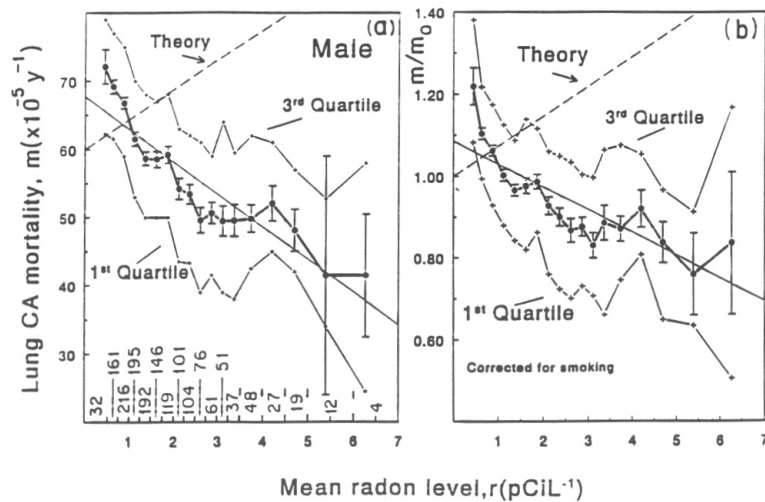
### 1. INTRODUCTION

#### 1.1. Background

In a 1995 paper (Cohen 1995), data on lung cancer mortality rates,  $m$ , in 1601 U.S. counties were compared with radon exposure,  $r$ , in those counties. These data are displayed in the plot in Fig. 1a, which is explained in the caption. It is evident from that figure that there is a strong and statistically indisputable tendency for lung cancer rates to decrease as radon exposures increase, which is contrary to initial expectations from the fact that radon causes lung cancer.

These data are used to test the validity of the linear-no threshold theory (hereafter, LNT) in the low dose region. This test is very different from other tests of LNT utilizing case-control studies (Lubin and Boice 1997), which are really designed to determine a risk vs dose relationship

Address correspondence to Bernard L. Cohen, 100 Allen Hall, University of Pittsburgh, Pittsburgh, PA 15260. E-mail: blc@pitt.edu. Telephone: (412)624-9245. Fax: (412)624-9163.



**FIGURE 1** Lung cancer mortality rates before (Fig. 1a) and after (Fig. 1b) correction for smoking prevalence vs average radon levels in homes, for 1601 U.S. counties. Data points shown are the average of ordinates for all counties within the range of  $r$ -values shown on the base-line of Fig. 1a; the number of counties within that range is also shown there. Error bars are one standard deviation of the mean, and the first and third quartiles of the distributions are also shown. Theory lines are arbitrarily normalized lines increasing at a rate of +7.3% per pCi/L as predicted (after the smoking correction) by LNT. These figures are used only for presentation; all analyses, including the straight line fit to the data shown here, use the 1601 actual data points.

for individual persons. That obviously requires data on individuals, whereas we have only average data on groups of individuals, the populations of counties. Such data on groups are called “ecological data”.

As an example of the difficulty this represents, consider a situation where the risk has a sharp threshold at 50 units of dose. The average risk in the county then depends on the fraction of the population exposed to more than 50 units, which is not necessarily related to the average dose which might be about 5 units. Clearly, the average dose does *not* determine the average risk, and is therefore not useful for determining the risk vs dose relationship. To assume otherwise is called “the ecological fallacy”. However it is readily demonstrated mathematically that this particular problem does not arise if the risk is linearly related to the dose. In that special case, the average dose *does* determine the average risk. This is familiar to Health Physicists from the widely used paradigm from LNT that person-sieverts (man-rem) determines the number of deaths; person-sieverts divided by population is the average dose, and number of deaths divided by population is the average risk.

The procedure for testing LNT involves two basic steps. The first step is, assuming LNT to be valid, to transform the risk vs dose relationship for individuals mathematically into a relationship between ecological variables, and the second step is to test that relationship against observation.

The first step starts with the BEIR-IV formula (NAS 1988) for risk to an individual, based on LNT, and develops it mathematically, summing over all persons in the county. This and subsequent analyses were done separately for males and females, always leading to similar results, but for brevity here, we confine our attention (with a single exception) to males. The result of the mathematical development (Cohen 1995) for males (with  $m$  in units of deaths per year/100,000 population, and  $r$  in units of  $37 \text{ Bq/m}^3$  [pCi/L]) is

$$M = m / [9 + 99 S] = A + B r \quad \text{Eqn. (1)}$$

where  $S$  is the fraction of adult males that smoke cigarettes,  $A$  is close to 1.0,  $B = +7.3$  (in percent increase per  $37 \text{ Bq/m}^3$  [per pCi/L]), and  $M$  is defined by the equation on the left and may be thought of as lung cancer rate corrected for smoking. The data thus corrected for smoking are shown in Fig. 1b.

Eqn.(1) is a relationship between ecological variables— $m$ ,  $r$ , and  $S$ —and hence it accomplishes our first step. Since it is derived mathematically from the LNT relationship between variables for individuals, if the latter is valid, Eqn.(1) must be valid and can be used as a test for the validity of LNT. This use of a mathematically derived formula to verify the theory from which it is derived is a time honored procedure in science. Newton’s famous theory relating force acting on an object, its mass, and its acceleration,  $\mathbf{F} = \mathbf{m} \mathbf{a}$ , was not directly tested for centuries since acceleration could not be directly measured; it was rather used to mathematically derive the distance traveled by the object vs time, which was measured to test the theory.

The fact that Eqn.(1), a relationship between ecological variables— $m$ ,  $r$ , and  $S$ —is being used to test LNT represents a radical departure from previous tests. This has far reaching consequences. The principal previous tests have used case-control studies for individuals, which require extensive information on these individuals. But in our approach, no such information is required unless it can be shown that it might affect the relationship between  $m$ ,  $r$ , and  $S$ . The difference this makes will be illustrated through the rest of this paper.

It is apparent from Fig. 1b that there is a huge discrepancy between the LNT prediction,  $B = +7.3$ , and the fit of Eqn.(1) to the observed data which gives  $B = -7.3 \pm 0.56$ , a discrepancy of 26 standard deviations. We refer to this as “our discrepancy”. The Scientific Method requires that, if a theory makes predictions that are discrepant with observations and if no plausible explanation can be found for that discrepancy, the theory is invalid. If LNT is to survive the test, we must therefore find a plausible explanation for our discrepancy. The principal purpose of this paper is to describe the search for such an explanation.

Before proceeding, it is important to understand that Fig. 1 should not be interpreted to be a dose-response relationship between radon exposure and lung cancer. As explained above, to do so would be falling into the trap of “the ecological fallacy”. There are only two logical alternatives to consider: (1) LNT is valid in which case a plausible explanation must be found for our discrepancy, or (2) LNT is not valid, in which case we cannot use these data to determine a dose-response relationship.

### 1.2. Confounding Factors (CF)

It is not unexpected that factors other than smoking and radon exposure can affect the risk for lung cancer. In principle, these should be included in the BEIR-IV formula for risk to individuals that we start with, carried through the mathematical development, and end up represented in Eqn. (1). This would be a completely unmanageable process, but failure to carry it out does not mean that the problem is unmanageable. Analogous situations arise universally throughout science. Few, if any, formulas used by scientists are absolutely exact and complete, not even Newton’s formula  $\mathbf{F} = m \mathbf{a}$ , but it is conventional to develop them mathematically to yield other formulas that are useful. In an elementary Physics course and in several Engineering courses, Newton’s formula is mathematically developed into dozens of useful formulas, involving trajectory of motion, kinetic and potential energy, momentum, torque, pressure, etc. In all fields of science, formulas simplified by neglecting less important terms are developed mathematically to derive other useful formulas. The neglected terms can be treated by various approximation methods, or simply ignored with a recognition that the results are subject to some uncertainty.

In seeking an explanation for our discrepancy, we must investigate the effects of variable factors that might, in principle, be included in a complete treatment of the lung cancer vs radon relationship. If one of these variables does indeed contribute to the lung cancer risk, and if, for some reason, it is correlated with radon levels, it would affect the relationship evident in Fig. 1. That variable would then be said to confound the relationship between  $M$  and  $r$ , and would be called a confounding factor (CF). As an illustrative hypothetical example, suppose that ozone levels in the atmosphere irritate the lungs and thereby cause lung cancer, and suppose that through some unknown process ozone scavenges radon out of the air, reducing radon levels. Then counties with high ozone levels would tend to have high lung cancer rates and low radon levels, and vice versa for counties with low ozone levels. The variations in ozone levels among U.S. counties would then, in itself, cause a negative slope,  $B$ , for the data in Fig. 1. This could possibly explain our discrepancy, indicating that county ozone levels would be an important CF.

A lengthy list of potential CFs could be drawn up, and each of these should be investigated before a judgment can be made on the validity of LNT. It is this process that we now describe. We begin by considering smoking prevalence to be known so that M and r have known values for each county, and later, in Section 4, we come back to consider potential confounding by smoking variables.

## 2. STRATIFICATION

### 2.1. The Stratification Method

A straightforward way to check on whether a particular factor, the value of which is known for each county, is a CF is to consider only counties for which that factor has the same value, leaving no possibility for it to confound. The practical manifestation of this procedure is to stratify the complete data file into many sub-files on the basis of the factor being investigated. As an example for which direct data are available, consider population density (PD) which might affect lung cancer rates through behavioral patterns and medical services, and might affect radon levels through house construction characteristics. In Table 1, the results are shown for stratifying our entire 1601 county data file into 10 deciles (sub-files) of 160 counties each on the basis of PD; the data in each decile are fitted to Eqn. (1) to obtain a completely independent value of the slope B of the M vs r regression. Table 1 includes data for females as an example of the general similarity of results for the two sexes, but data for

**TABLE 1** Treatment of “County Population Density” (PD) as a confounding factor by the stratification method. Results are for B using single regression of M on r as in Eq. (1), and multivariate (double) regression of M on r and PD, fitting the data to

$$M = A + B r + E PD$$

where E is a fitting parameter. Bottom line gives the averages of the columns above and the standard deviation of that average.

County Rank by PD	PD range (×100/sq.mi)	Single Regression		Double Regression	
		B-male	B-female	B-male	B-female
1–160	0.003–0.094	–3.7	–6.6	–3.7	–6.4
161–320	0.095–0.22	–8.0	–7.8	–8.0	–7.9
321–480	0.22–0.35	–7.0	–8.5	–7.0	–8.5
481–640	0.35–0.50	–6.4	–9.7	–6.4	–9.8
641–800	0.50–0.67	–8.9	–8.7	–8.9	–8.7
801–960	0.67–0.92	–4.3	–4.4	–4.3	–4.4
961–1120	0.93–1.29	–9.2	–6.0	–9.3	–6.0
1121–1280	1.30–2.05	–5.9	–8.1	–5.9	–8.1
1281–1440	2.05–4.11	–0.5	–2.7	–0.5	–2.8
1441–1601	4.12–671.8	–4.5	–7.4	–3.9	–6.2
Average ± Std.Dev.		–5.8±2.7	–7.0±2.1	–5.8±2.7	–6.9±2.1

females will be omitted for brevity in all further discussions. From the second column of Table 1, we see that in each stratum (except the last) the values of PD are very similar, much more similar than in the data for all U.S. counties, so any confounding by PD is greatly reduced. This is reiterated in Table 1 by including results of a multivariate regression of M on r and PD, noting that B-values (the coefficients of r in the regression) from single and multivariate regression are essentially the same. Note that the values of B are all negative and generally of the same magnitude as the value for the entire 1601 county data set. The differences between their average and the values for the entire data set,  $B = -7.3$  for males and  $B = -8.3$  for females, are well within the standard deviation of the averaging process. More importantly, there is no evident trend for B to increase or decrease monotonically with increasing PD; there is little difference in B-values if we consider only counties with the largest PD, or if we consider only counties with the smallest PD, or if we consider only counties with average PD, etc. These facts clearly indicate that confounding by PD is of little relevance in explaining our discrepancy.

## 2.2. Comparison of Stratification with Multivariate Regression

Since stratification is a somewhat laborious process, one might ask why not simply do a multivariate regression of M on r and the CF, and accept the coefficient of r in this regression as the value of B corrected for confounding? One obvious weakness of multivariate regression is that it assumes that the relationship of M to the CF is a linear one, which may not be true. But here we offer a treatment which demonstrates a more subtle weakness.

We begin by recognizing the fact that the only way a confounding factor, X, can affect the value of B derived from fitting data with Eqn.(1) is by systematically causing counties with low M to have high (or low) r, and vice versa. This would be evidenced by the rankings of counties in our data file by M,  $R(M)$ , which has a value ranging from 1 to 1601 for each county, and (for our case) the inverse rankings of these counties by r,  $R(r)$ , both being highly correlated, for unrelated reasons, with the rankings of counties by X,  $R(X)$ . We refer to these correlations by ranking (also known as Spearman's  $\rho$ ) as  $CoRR(X,M)$  and  $CoRR(X,r)$  respectively; that is, in a notation where  $Corr(a,b)$  denotes the coefficient of correlation (Pearson product moment) between a and b,

$$\begin{aligned} CoRR(X,M) &= Corr[R(X), R(M)] \\ CoRR(X,r) &= Corr[R(X), R(r)] \end{aligned}$$

Both  $CoRR(X,M)$  and  $CoRR(X,r)$  must be large for unrelated reasons if X is to be an important confounding factor.

Using this background, we now address the issue raised at the beginning of this section, on the use of multivariate regression. Let us assume that there is a confounding factor, X, that is causally related to M but has *no* causal or other direct relationship to r. It thus cannot confound the relationship between M and r, and therefore should have no effect on the value of B. But it will necessarily have a correlation with r because of the correlation of M with r evident in Figure 1.

To study this effect quantitatively, we take

$$R(X) = s R(M) + (1 - s) R\text{-random}$$

where R-random is a random rearrangement of rankings and s can be varied between 0 and 1.0 to get various CoRR(X,M). Some results for CoRR(X,r) and B, defined here as the coefficient of r in the multivariate regression of M on r and R(X), follow:

CoRR(X,M)	0	0.30	0.44	0.59	0.73
CoRR(X,r)	0	-0.14	-0.18	-0.23	-0.28
B	-7.3	-6.7	-6.0	-4.9	-3.7

We see that, even though X is not a confounding factor in the relationship between M and r because it has no causal or other direct relationship with r, its use in multivariate regression still has an appreciable effect on the resulting value of B. This demonstrates why the method of stratification is preferable to multivariate regression in assessing the effects of confounding factors; since with stratification, the value of the CF is essentially the same for all data in each regression, it cannot affect that regression.

The demonstration here does not mean that multivariate regression is a useless tool. But it indicates that it should be used with some caution.

### 2.3. Problem: Average Value of a CF May Not Represent Its Confounding Effects

The stratification procedure may not eliminate effects of a confounding relationship because the average value of a CF does not necessarily represent its confounding effects. For example, average annual income may not represent the confounding effects of monetary income because its confounding effects may depend on the fraction of the population that is very poor, or very rich. To cover this problem, we consider separately as CF the fraction of the population with income <\$5000, \$5000 to \$10,000, . . . . . , >\$150,000 (10 brackets in all), plus various combinations of adjacent brackets (Cohen 2000a). Since none of these has a confounding effect, we may conclude that any aspect of annual income, not just *average* annual income, can be excluded as an important CF.



As a related but different type example, a person's age is an explicit factor in the BEIR-IV formula for risk vs dose to an individual, and it is carried through the mathematical process of deriving Eqn.(1) by showing that variations in age distributions do not have appreciable effects on that equation (Appendix A of Cohen 1995). But as an extended treatment of that issue, we consider as CF the fraction of the population in age groups <1 year, 1–2 years, . . . . . , >85 years (31 groups in all) plus various combinations of adjacent brackets. Since none of these has an important confounding effect, we may conclude that age distribution can be excluded as a CF explaining our discrepancy.

Perhaps our conclusions here, that annual income and age distribution are not plausible confounders, is not justified. To consider this possibility, we need a suggestion for specific not-improbable dependences on these for individuals that would not be reflected in the ecological CF specified above. Despite frequent efforts to devise or solicit such a suggestion, none has materialized.

#### **2.4. Stratification on Geography and Environmental Factors**

The very wide differences in average radon levels in counties, evident in Fig. 1, indicates that geography is an important factor in determining radon levels. Since different geographic regions have many other different characteristics that may affect lung cancer rates—climate or ethnicity of the population are plausible examples—geography is a potentially important CF. This can be investigated with a method akin to stratification by dividing the nation into sections and doing a separate analysis for each section.

This was done using Bureau of Census Regions and Divisions, and using individual states (Cohen 2000a). If we take the average value of B to represent a corrected true value, the results are as follows: for the 4 national regions,  $B = -5.2$ ; for the 8 national divisions,  $B = -4.1$ ; for the 33 individual states plus 4 combinations of contiguous states (combined so as to get at least 20 counties in each data file),  $B = -5.0$ . This may be interpreted as indicating that confounding by geography changes the value of B from  $-7.3$  to about  $-5.0$ , but this is still a long way from explaining our discrepancy with the LNT value,  $B = +7.3$ .

Stratification on environmental factors like altitude, temperature, precipitation, etc was treated in Section K of (Cohen 1995) and found not to affect the results. The behavior shown in Fig. 1 is found if we consider only the warmest areas or if we consider only the coolest areas, if we consider only the wettest or only the driest, etc.

#### **2.5. Screening Candidates for Stratification Studies**

Well over 100 potential CF have been treated by the stratification method with no progress in resolving our discrepancy. But stratification



is a tedious process and the number of potential confounding factors is very large. A more rapid screening procedure is desirable. Moreover, there are numerous potential CF for which no data are available—ozone levels discussed in Section 1.2 above is an example—but they cannot simply be ignored. The solution to these problems is treated in Section 3 below.

### 3. OTHER CONFOUNDING FACTOR ISSUES

#### 3.1. Plausibility Requirements on Confounding Factors

For subsequent analyses, it is important to understand quantitatively what is required for a confounding factor to influence our results. As pointed out in Section 2.2, the only situation in which a confounding factor X can affect the value of B derived from fitting data with Eqn (1) is where both CoRR(X,M) and CoRR(X,r) are large for definite but unrelated reasons. The most effective R(X), was found to be

$$R_0(X) = \text{Ranking of } \{0.5 R(M) + 0.5 R(r)\}$$

which, applied to our data file, gives CoRR(X,r) = CoRR(X,M) = 0.82. Lesser correlations can be generated by taking

$$R(X) = \text{Ranking of } \{p R_0(X) + (1 - p) R\text{-random}\}$$

where R-random is a random rearrangement of the rankings, that is of integers between 1 and 1601, and p is varied between 0 and 1.0 to obtain varying CoRR(X,M) and CoRR(X,r). For each value of p we utilize stratification on R(X) into quintiles, fitting the data in each stratum to Eqn (1) to obtain a B-value for that stratum, and averaging the B-values from the five strata to obtain a B-value for the entire data set. The results for three different sets of R-random (generated by the MINITAB statistical pack-

**TABLE 2** B values obtained if a confounding factor, X, has various correlations by ranking with M and r, CoRR(X,M) and CoRR(X,r). The three sets of results are for three different R-random. The first B-value is from stratification into quintiles, and value in parenthesis is the coefficient of r in a multivariate regression of M on r and R(X).

CoRR (X,r)	CoRR (X,M)	B	CoRR (X,r)	CoRR (X,M)	B	CoRR (X,r)	CoRR (X,M)	B
0.09	0.09	-7.2(-7.2)	0.07	0.12	-7.2(-7.2)	0.08	0.08	-7.2(-7.2)
0.18	0.18	-7.0(-6.9)	0.16	0.21	-6.9(-6.8)	0.16	0.17	-7.0(-6.9)
0.34	0.34	-5.2(-5.5)	0.32	0.37	-5.3(-5.4)	0.33	0.33	-5.1(-5.6)
0.53	0.53	-2.1(-2.2)	0.51	0.55	-2.3(-2.0)	0.52	0.52	-2.1(-2.3)
0.69	0.69	+3.2(4.0)	0.68	0.70	+2.9(+4.0)	0.69	0.69	+3.3(+3.9)
0.79	0.79	10.5(11.3)	0.78	0.79	10.4(11.1)	0.79	0.79	10.8(11.2)
0.81	0.81	13.7(14.4)	0.81	0.82	13.7(14.2)	0.81	0.81	13.8(14.3)

age) are listed in Table 2. The B-values in parentheses there are the coefficients of  $r$  from multivariate regression of  $M$  on  $r$  and  $R(X)$ . Note that the B-values obtained in these two very different ways are in good agreement. In view of our definition of  $R_0(X)$ ,  $\text{CoRR}(X,M)$  and  $\text{CoRR}(X,r)$  should be nearly the same, but since these depend somewhat on R-random, both are listed in Table 2.

We see from Table 2 that the different sets of R-random give consistent results and indicate that for a CF,  $X$ , to shift the value of  $B$  from its original value,  $B = -7.3$ , to the LNT prediction  $B = +7.3$ , requires  $\text{CoRR}(X,r)$  and  $\text{CoRR}(X,M)$  (or, according to further calculations, their average) to be about 0.75, and even to change the sign of  $B$  from  $-$  to  $+$ , accounting for half of our discrepancy, requires these correlations to be about 0.6.

The bottom rows of Table 2 show that a confounding factor can indeed drastically change the results of the study, as Lubin (1998) has demonstrated mathematically. But there is an unstated corollary to Lubin's mathematical demonstration—the required values of the CF must be plausible. The issue of plausibility must be addressed.

How plausible are the values of a CF leading to the correlations required here? The factors affecting radon exposure,  $r$ , are geology and house construction characteristics, while the factors affecting  $M$ , lung cancer rates corrected for smoking, are human behavioral and genetic characteristics, so it is very difficult to imagine a CF, other than geography which was treated in Sec. 2.4, that has a causal relationship with both of these very different type characteristics. It would seem that the most likely source of confounding is through socioeconomic variables, SEV. Our data base includes 530 SEV (Cohen 2000a). For these, the maximum (in absolute value)  $\text{CoRR}(\text{SEV},r)$  is 0.486; for only 13 of these 530 SEV is it  $>0.4$ , and for only 49 of them is it  $>0.3$ . The maximum  $\text{CoRR}(\text{SEV},M)$  is 0.39 and for only 13 of the 530 SEV is it  $>0.3$ . Calculations indicate that the relevant quantity is the average of the absolute values of  $\text{CoRR}(\text{SEV},r)$  and  $\text{CoRR}(\text{SEV},M)$ , which we call "Aver-CoRR" (in cases where the signs of the two are opposite). The maximum Aver-CoRR is 0.43, only one other is  $>0.4$ , only 11 are  $>0.35$ , and only 25 of the 530 SEV have Aver-CoRR  $>0.3$ . It thus seems implausible for any CF to have Aver-CoRR larger than 0.5, and it is surely very highly implausible for Aver-CoRR to approach 0.75, which is required to explain our discrepancy with the LNT prediction, or even 0.6 which is required to substantially reduce that discrepancy.

It should be noted in passing that all of the strong correlations cited above can be explained as arising from the urban-rural effect—urban people smoke more and have lower radon exposures than rural people. The effects of this on B-values were studied by stratification on county population, percent urban, etc in Section H of (Cohen 1995) and no effects on

B-values were found. The urban-rural effect within counties was treated in Section L of (Cohen 1995) and found not to affect the value of B.

The concept of plausibility of correlation introduced above is a very powerful one, available in this study because there are data on such a large number of potential CF, enough to draw meaningful conclusions about the distribution of their correlations. For example, it covers ozone as a CF in the hypothetical situation introduced in Sec 1.2. Ozone level in the atmosphere is related to urban vs rural factors, importance of manufacturing, prevalence of motor vehicles and highways, and other variables for which data are available and included in our analyses. We may thus conclude that ozone level is not an important CF even though data are not available on ozone levels in each county. Similarly, we may conclude that any factor which is related to socioeconomics may be excluded as a CF that might explain our discrepancy.

### **3.2. Confounding factors on the level of individuals**

There are potential CF on the level of individual people, rather on the level of groups of people like county populations, that might seem not to be represented by ecological variables, as required in our procedures. As a “far-out” example, one might think this applies if a man’s lung cancer risk depends on  $Y = [\text{the product of his annual income squared, and the number of siblings that he has, raised to the fourth power}]$ . But, in principle, counties could keep statistics on the values of  $Y$  in its population, and report averages,  $Y_{AV}$ . These would then be an ecological socioeconomic variable, and it would be reasonable to expect  $\text{CoRR}(Y_{AV}, r)$  to be in the same range as other  $\text{CoRR}(\text{SEV}, r)$ , which would not affect our results.

There is a substantial literature pointing out that CF on the level of individuals cannot be adequately represented in a case-control study by ecological variables (Greenland and Robins 1994, Morgenstern 1995, Stidley and Samet 1994, Lubin 1998). But before this can be interpreted as invalidating our study, it must be shown that such a CF can affect the relationship between  $m$ ,  $r$ , and  $S$ , which is the basis for our test of LNT through Eqn. (1). In trying to go through this process, I have found it to be inevitable that the effect under consideration can be represented by ecological variables; the effects considered below in Sec. 4.3, 4.4, and 4.5 are examples, as is the treatment of  $Y$  in the previous paragraph. I would be anxious to address any suggestions for effects in which this treatment fails.

### **3.3. Effects of Combinations of Confounding Factors**

Up to this point we have been considering CFs one at a time. Since any one may cause small changes in the value of B, is it possible that these

small changes can accumulate and thereby explain our discrepancy? We address that issue here.

The only way in which confounding factors can affect the results is if the rankings of counties by these factors is highly correlated for unrelated reasons with  $R(M)$  and  $R(r)$ . But from the treatment of CF by stratification, it is clear that only one set of such rankings can enter into the determination of  $B$  through its correlations with  $R(M)$  and  $R(r)$ . This could be an equivalent set,  $R(E)$ , based on all relevant CFs. It is important to recall here that  $\text{CoRR}(X,r)$  refers to the *inverse* ranking of  $X$  vs  $r$ ; the correlations considered in Table 2 require that  $X$  be correlated with  $M$  and  $r$  in opposite directions. There can just as easily be confounding factors  $X$  that are correlated in the same direction with  $M$  and  $r$  in which case the effect is to make the  $B$ -value more negative, increasing the discrepancy with the LNT prediction. For example, with  $\text{CoRR}(X,r) = \text{CoRR}(X,M) = 0.75$  with correlations in opposite directions, analysis gives  $B = +7.7$ , but for these correlations in the same direction, the result is  $B = -16.2$  which is a much larger discrepancy with LNT than is found without confounding,  $B = -7.3$ . Thus  $R(E)$  is just as likely to increase our discrepancy as to reduce it.

But improbable as it is, let us consider the worst case, where all effective CF have correlations with  $r$  and  $M$  in opposite directions. If we define  $R(1), R(2), R(3), \dots$  as the rankings of the most important CF, the second most important CF, the third most important CF,  $\dots$ , a first approximation to  $R(E)$  is  $R(1)$ . Any improvement to  $R(E)$  by making changes to include  $R(2), R(3)$ , etc will decrease the effectiveness of  $R(1)$ , and hence will tend not to be a major improvement. Thus the effect of combinations of confounding factors would not be much greater than the effect of the single most important confounding factor.

As a test of this conclusion, the confounding effects of the 530 SEV in our data base were investigated by multivariate regression of  $M$  on  $r$  and groups of SEV. If we use multivariate regression of  $M$  on  $r$  and the single SEV with the largest Aver-CoRR, the slope  $B$  is changed from its original value,  $B = -7.3$ , to  $B = -4.8$ . If we expand the multivariate regression to include 12 variables,  $r$  plus the 11 SEV with the largest Aver-CoRR, the result is  $B = -4.4$ , only a slight change from the effect of the single most important CF.

Since the combination of all CFs cannot have much more effect than that of the single most important CF, and we have shown that it is very highly implausible for a single CF to explain our discrepancy, we conclude that it is also very highly implausible for a combination of CF to affect the results.

Accepting the result that confounding may change  $B$  from  $-7.3$  to  $-4.4$  reduces our discrepancy with the LNT prediction,  $B = +7.3$ , by only 20%, not an appreciable improvement. Moreover, it was shown in Section

I of (Cohen 1995), using an approach similar to that in Section 2.2, that use of multivariate regression substantially over-estimates the effects of confounding.

#### **4. CONFOUNDING BY SMOKING-RELATED VARIABLES**

##### **4.1. Relative Importance of Smoking and Radon Exposure**

One might think that smoking is such a dominant cause of lung cancer that its effects can easily mask the effects of radon. To address this, we estimate the relative importance of these two factors in determining lung cancer rates by use of BEIR-IV. The width of the distribution of S-values for U.S. counties, as measured by its standard deviation, is 13.3% of the mean, which, according to BEIR-IV, would cause a difference in lung cancer rates of 11.3%. The width of the distribution of average radon levels is 58% of the mean, which, according to BEIR-IV, would cause a difference in lung cancer rates of 6.6%. Thus the importance of smoking for determining variations in lung cancer rates in U.S. counties is less than twice (that is,  $11.3 / 6.6$ ) that of radon exposure.

But even more important for our purposes is the fact that smoking prevalence,  $S$ , can only influence our results to the extent that it is correlated with radon levels,  $r$ . Thus we are facing a straightforward quantitative question: How strong an  $S$ - $r$  correlation is needed to affect our results? That question is addressed in the next section.

##### **4.2. Uncertainties in Smoking Prevalence, $S$**

Smoking prevalence,  $S$ , has a very special place in our analysis due to its explicit inclusion in Eqn. (1). Since  $S$  is involved in the equation to be fitted, the distribution of S-values, not just its ranking for various counties, affects results for  $B$ . Three very different sources of data were used to determine S-values for counties—(1) a Bureau of Census survey for States with an adjustment for urban vs rural differences among the counties in each state, (2) state cigarette sales tax collections with a similar adjustment, and (3) lung cancer rates for counties with similar radon levels. Each of these utilized separately gives essentially the same results as those based on combining the three. Nevertheless, the uncertainty in S-values was still a matter of some concern which was addressed by studying the correlations between  $S$  and  $r$  required to explain our discrepancy.

As an initial approach, values of the best estimated S-values are maintained but these S-values are reassigned to counties so as to give  $\text{CoRR}(S.r) = -1.0$ ; that is, the county with the lowest  $r$  was assigned the highest  $S$ , the county with the next lowest  $r$  was assigned the next highest  $S$ , and so forth thru our 1601 counties, ending with the county with the highest  $r$  assigned the lowest  $S$ . Even with this perfect inverse correlation,

which completely violates any considerations of plausibility, B is only reduced from its original value,  $-7.3$ , to zero, still leaving half of our discrepancy with the LNT prediction,  $B = +7.3$ , unexplained.

Going still further, the effects are increased if the distribution of S-values is wider. The maximum not implausible width for the distribution of S-values is the width of the lung cancer mortality rate (m) distribution, since other factors influence m in ways that, statistically, would increase that width. With this increased S-distribution width, centered on the well established national average for S, S-values are reassigned to each county to give  $\text{CoRR}(S,r) = -1.0$ ; we call this S-perfect. At the other extreme, these same S-values are randomly assigned to each county to obtain S-random. Calculations are then done with

$$S = q \text{ S-perfect} + (1 - q) \text{ S-random} \quad (2)$$

where q is various numbers between 0 and 1.0 chosen to obtain various  $\text{CoRR}(S,r)$  and coefficients of correlation with r (*not* correlations by rank),  $\text{Corr}(S,r)$ . The results for three different sets of S-random are shown in Table 3. We see there that the  $\text{Corr}(S,r)$  required to change B to the LNT prediction,  $B = +7.3$ , is about 0.9, and just to reduce B down to zero, eliminating half of the discrepancy, is about 0.62, even with this substantially increased width of the S-distribution.

How plausible are these required  $\text{Corr}(S,r)$ ? The most probable source of a correlation between S and r is through socioeconomic variables, SEV. It therefore seems reasonable to assume that  $\text{Corr}(S,r)$  should be in the same range as  $\text{Corr}(\text{SEV},r)$  for other SEVs. In our data base of 530 potential confounding SEV, the largest  $\text{Corr}(\text{SEV},r)$  is 0.45, only 7 of the 530 are  $>0.4$ , and only 15 are  $>0.35$ . It thus seems reasonable to assume that values of  $\text{CoRR}(S,r)$  larger than 0.5 are implausible. It is surely highly implausible for  $\text{Corr}(S,r)$  to approach the values,  $0.62 - 0.90$ , required to help explain our discrepancy, even if we accept the almost

**TABLE 3** B-values obtained if smoking prevalence, S, has various  $\text{Corr}(S,r)$ , assuming the maximum plausible width for the S-distribution. The three sets of results are for three different S-ran-

Corr (S,r)	CoRR (S,r)	B	Corr (S,r)	CoRR (S,r)	B	Corr (S,r)	CoRR (S,r)	B
-0.17	-0.17	-7.1	-0.24	-0.23	-5.5	-0.23	-0.22	-6.0
-0.33	-0.32	-4.7	-0.39	-0.37	-3.2	-0.38	-0.36	-3.7
-0.41	-0.39	-3.5	-0.47	-0.45	-2.0	-0.45	-0.43	-2.6
-0.49	-0.47	-2.3	-0.54	-0.52	-0.9	-0.53	-0.51	-1.4
-0.57	-0.55	-1.1	-0.62	-0.60	+0.3	-0.60	-0.58	-0.2
-0.65	-0.63	+0.1	-0.68	-0.66	+1.4	-0.68	-0.66	+0.9
-0.78	-0.75	+2.7	-0.81	-0.79	+3.8	-0.80	-0.78	+3.3
-0.88	-0.85	+5.5	-0.89	-0.86	+6.4	-0.88	-0.85	+6.0
-0.93	-0.90	+8.6	-0.93	-0.90	+9.3	-0.93	-0.90	+8.9

implausibly increased width of the distribution of S-values which ignores our three independent sources of data.

### 4.3. Different r for Smokers and Non-smokers

Another type problem arises if there is a systematic difference in average radon exposures for smokers,  $r_s$ , and non-smokers,  $r_n$  (Cohen 1998a). Since smokers are 12 times more at relative risk from radon than non-smokers (NAS 1988), the effective radon level,  $r_e$ , for the county as a whole for causing lung cancer is

$$r_e = [12 S r_s + (1 - S) r_n] / [12 S + (1 - S)]$$

where the two terms in the numerator are the weightings for radon exposure to smokers and non-smokers, and the denominator is the sum of these weightings. This differs from the measured average radon level,  $r$ ,

$$r = S r_s + (1 - S) r_n$$

If we define  $x = r_s / r_n$ , the relationship between the effective and measured radon levels is converted by algebra to

$$r_e = r (12 S x + 1 - S) / [(x S + 1 - S) (11 S + 1)]$$

We then use  $r_e$  instead of  $r$  in fitting the data to determine values of B. In doing this, the parameters that may be varied are the average value of  $x$  ( $x$ -average), the width of the distribution of  $x$ -values, and  $\text{Corr}(x,r)$ .

It has been found (Cohen 1991) that the national average for  $x$  is 0.9, but we give some results for other values of  $x$ -average. For the 52 of the 54 SEV considered in (Cohen 1995) that are not proportional to the county population, the average width of distributions is 26 % of their mean, and for only one of the 52 is it above 50%—55% for “percent of income from government” which is an understandable special case. On this basis, we consider distributions of  $x$ -values to have width 57% of the mean, which severely stretches the limits of plausibility, and 28% of the mean which is in the region of reasonable plausibility.

Some results are listed in Table 4. The first five entries explore the effect of  $x$ -average using assumptions about the other factors most favorable for explaining the discrepancy. The remaining entries use the known value of  $x$ -average and explore the effects of the width of the distribution and of  $\text{Corr}(x,r)$ . The results in Table 4 indicate that it is highly implausible for systematic differences between radon exposures to smokers and non-smokers to change B from  $-7.3$  to less than about  $-5.5$ , still a very long way from the LNT prediction,  $B = +7.3$ .



**TABLE 4** Effects of difference in radon exposure for smokers and non-smokers, with x = smoker/non-smoker exposures in each county. Table gives value of B for various choices of the width of distribution of x-values (SD as fraction of mean) and Corr(x,r).

x-average	SD of x	Corr(x,r)	B
0.8	0.57	1.0	-4.9
0.9	0.57	1.0	-4.8
1.0	0.57	1.0	-4.7
1.2	0.57	1.0	-4.5
1.5	0.57	1.0	-4.3
0.9	0.57	0	-6.5
0.9	0.57	0.4	-5.9
0.9	0.57	0.7	-5.5
0.9	0.57	1.0	-4.8
0.9	0.28	0	-7.3
0.9	0.28	0.4	-6.7
0.9	0.28	1.0	-5.6

#### 4.4. Variations in Intensity of Smoking

The BEIR-IV formula for cancer risk to an individual, which was the starting point for our test of LNT, considers only the distinction between smokers and non-smokers, with no consideration of intensity of smoking. Therefore, that factor is not represented in Eqn. (1) which is derived from the BEIR-IV formula. But the BEIR-VI Report (NAS 1999) suggested that Eqn. (1) is deficient in that it ignores intensity of smoking, and proposes that this be treated by dividing smokers into two categories, 2 pack/day and 1 pack/day. To study this (Cohen 2000b), we define

k = ratio of 2 pack/day to 1 pack/day smokers in a county

f = ratio of lung cancer risk for 2 pack/day to 1 pack/day

Analysis of available data indicates the plausible values most favorable for the BEIR-VI suggestion are f = 2.0 and national average for k = 0.4. Using these converts Eqn. (1) to

$$M = m / [9 - 9S + 84 S \{(1 + 2 k)/((1 + k))\}] = (A + B r) \quad (3)$$

Different distributions of k-values were tried but the most promising was a level distribution between 0 and 0.8, to be consistent with the national average of 0.4. We assign k-values to counties so as to define k-perfect as assignments for which CoRR(k,r) = 1.0, and k-random as one where k-values are assigned randomly. We then generate k-values to be used in fitting Eq. (3) as

$$k = g \text{ k-perfect} + (1 - g) \text{ k-random}$$

where  $g$  is given various values between 0 and 1.0 to obtain different  $\text{Corr}(k,r)$ . The results are

$\text{Corr}(k,r)$	0	-0.37	-0.52	-0.78	-0.91	-0.93
B	-7.6	-5.0	-4.2	-2.5	-0.7	+1.4

In view of our previous discussion on plausibility of correlations, it seems reasonable to assume that an absolute value of  $\text{Corr}(k,r) > 0.5$  is highly implausible. It is therefore clear that including intensity of smoking as a confounder can do little to reduce our discrepancy with the LNT value,  $B = +7.3$ . Of course, there is no reason why  $\text{CoRR}(k,r)$  should not be negative rather than positive, in which case the discrepancy with LNT would be increased, so the above table gives an unbalanced view, emphasizing things that may reduce the discrepancy with LNT.

(Cohen 2000b) also considers possible correlations with  $r$  for both  $S$  and  $k$ , using the above method. For cases where  $\text{Corr}(k,r) = \text{Corr}(S,r)$ , as these vary from zero to  $-0.8$ ,  $B$  increases roughly linearly from  $-10.0$  to  $+1.3$ ; for example, for  $\text{Corr}(k,r) = \text{Corr}(S,r) = 0.4$ ,  $B = -4.3$ . Again it is apparent that plausible values of these correlations can do little to bring  $B$  close to the LNT prediction,  $B = +7.3$ .

#### 4.5. Combinations of Confounding by Smoking and Other Factors

To go further, we consider the effects of uncertainties in our  $S$ -values combined with an unknown confounding factor,  $X$ . Starting with our best estimates of  $S$ -values in a pool, we reassign  $S$ -values from this pool utilizing Eqn (2) to generate sets of  $S$ -values with various  $\text{Corr}(S,r)$ . For each of these sets of  $S$ -values, we determine corresponding sets of  $M$ -values, uti-

TABLE 5 B-values from combined effects of various  $\text{CoRR}(S,r)$  and  $\text{CoRR}(X,r)$

$\text{CoRR}(X,r)$	$\text{CoRR}(S,r)$				
	-0.69	-0.53	-0.37	-0.23	0.00
-0.65	6.5	5.5	4.4	3.3	1.5
-0.60	4.4	3.3	2.0	1.0	-1.2
-0.55	2.4	1.1	0.2	-1.0	-2.9
-0.50	1.0	0.0	-1.0	-1.7	-4.5
-0.45	-0.4	-1.6	-2.5	-3.6	-5.5
-0.40	-1.4	-2.4	-3.5	-4.5	-6.8
-0.35	-2.2	-3.2	-4.3	-5.4	-7.6
-0.30	-3.0	-4.0	-5.1	-6.1	-8.3
-0.25	-3.5	-4.5	-5.6	-6.6	-8.7
-0.20	-4.0	-5.1	-6.2	-7.0	-9.2
0.0	-4.8	-5.9	-6.9	-8.0	-10.0

lizing the left side of Eqn (1). For each of these sets of M-values, we go through the analysis described at the beginning of Sec. 3.1, to obtain sets of R(X) with various CoRR(X,r); we then determine values of B as the coefficient of r in multivariate regression of M on r and R(X). This gives tables of B-values for various combinations of Corr(S,r) and CoRR(X,r). By interpolating from these tables, we derive Table 5 which shows values of B obtained for various combinations of CoRR(S,r) and CoRR(X,r).

Applying our plausibility limit of 0.5 to both CoRR(S,r) and CoRR(X,r) simultaneously, which is far less plausible than applying it to only one of the two, and interpolating in Table 5, we obtain  $B = -0.2$ . This is still strongly discrepant with the LNT prediction,  $B = +7.3$ .

## 5. OTHER PROBLEMS WITH CONFOUNDING

### 5.1. Urban vs Rural Differences

In extensive studies (Cohen 1991) of how radon levels vary with socioeconomic factors, house characteristics, geography, etc), it was found that rural houses average about 25% higher radon levels than urban houses, whereas urban males smoke about 25% more frequently than rural males. This problem was treated in Section L of (Cohen 1995) using a model with the above percentages as parameters, by modifying the derivation of Eqn.(1) to consider not just the two categories, smokers and non-smokers, but four categories, urban and rural smokers and urban and rural non-smokers, each category having its own percentage of the population, lung cancer rate, and average radon level. These are related by the percent of the population that lives in urban areas, a known quantity for each county, and m, r, and S for the county. It was found that the changes in B caused by various plausible values of the parameters was only a few percent.

### 5.2. Differences between Radon Exposure and Measured Radon Levels

There have been suggestions that effective radon exposure, r-effective, may not be the same as the measured radon level in the home, r-measured; for example, time spent in the home may vary, or exposures outside the home may be important. We represent this (Cohen 1998b) as

$$r\text{-effective} = (1 + f) r\text{-measured}$$

The properties of f that can affect our results are the width of the distribution of f-values among the counties, and the correlation between f and r, Corr(f,r). We take the distribution to be uniform between  $-w$  and  $+w$ , where w is a parameter to be used in calculations. To test the maximal effects of

**TABLE 6** Slopes B from regression of M on r-effective for various values of w and Corr(f,r). The last column is the standard deviation in the determination of B.

w	Corr(f,r)	B	SD(B)
0		-7.3	0.56
0.2	0	-7.0	0.53
0.5	0	-5.5	0.47
0.8	0	-4.0	0.41
1.0	0	-3.2	0.36
0.1	+1.0	-6.4	0.48
0.3	+1.0	-5.2	0.39
0.5	+1.0	-4.3	0.33
0.7	+1.0	-3.7	0.28
0.1	-1.0	-8.5	0.62
0.3	-1.0	-12	0.9
0.5	-1.0	-21	1.4

correlations between f and r, values of f can be assigned to counties such that  $CoRR(f,r) = 1.0$ , which is equivalent to  $Corr(f,r) = 1.0$ . To check on the effects of no correlation, values of f can be assigned randomly.

The values of the slope, B, from regression of M on r-effective, are shown in Table 6, along with standard deviations in the B-values derived from the regression analysis. We see from Table 6 that no values of w or Corr(f,r) can help substantially in explaining our discrepancy with LNT predictions. In fact, when these factors reduce the negative value of B, they also reduce the standard deviation, so the number of standard deviations by which B differs from the LNT prediction,  $B = +7.3$ , is not reduced.

## 6. COMPARISON WITH CASE-CONTROL STUDIES

Since case-control study practitioners usually deal with risk to individuals, their studies require data on CFs for individual persons, obtained by questioning each involved individual or a close relative or acquaintance. This information is of key importance in their studies. For example, if annual income is an important element, the annual income of each individual must be connected to his cancer or lack of cancer. They may sometimes use ecological data for crude estimates; in the above example, they might assume the annual income of each individual to be the average income in his section of the city. But they clearly recognize this to be an inferior procedure, and label studies that depend on such procedures as “qualitative”, useful only for suggesting “analytical” studies that avoid them. Non-epidemiologists have frequently used ecological data unjustifiably to imply risk vs dose relationships for individuals. It is easy to show

how any paper depending on ecological data for such a purpose can give false results. It is certainly easy to understand that epidemiologists are instinctively “turned off” by use of ecological data.

However, our test of LNT is based on Eqn.(1) which is a relationship between ecological variables. Analyses are therefore straightforward for ecological CFs. Fortunately, every potential confounding relationship that seems plausible to me, or that has been suggested as being plausible by others, can be represented by ecological variables.

As an example of the difference between our approach and that of a case-control study, consider a hypothetical situation in which people of a certain ethnicity, call it Ethnicity-A, may have a high risk for lung cancer. In our approach, the relevant variable for determining the county lung cancer rate is the fraction of the population of Ethnicity-A, an ecological variable; there is no need to know which individuals in the county are of that ethnicity. If we were trying to find out whether people of ethnicity-A have an excess cancer risk, it would not be sufficient to find high lung cancer rates in counties with large fractions of their citizenry of ethnicity-A. We would have to know whether it is the people of ethnicity-A in those counties who had excess lung cancer. But in our study, we are simply testing the consequences of the hypothesis that people of ethnicity-A might have an excess cancer risk, which might cause counties with large populations of ethnicity-A to have high lung cancer rates, which could affect our results (if people of ethnicity-A have systematically low radon exposure). The fact that we find our results unaffected by fraction of the citizenry of ethnicity-A does not disprove the hypothesis that people of ethnicity-A may have high cancer risk, but that is irrelevant to our purpose which is to find CFs that *do* affect our results.

At least two papers (Lagarde and Pershagen 1999, Darby, Deo, and Doll 2001) have pointed out cases where the relationship between lung cancer and radon exposure derived from a study of individuals gives results different from an ecological study based on the same data. But these ecological studies involved *no* treatment of confounding factors, and the difference between results from these and from the individual level studies is easily explained by recognizable confounding factors. That is certainly not the case in our study which involves very extensive consideration of possible confounding factors.

It is frequently implied that our study is inferior to case-control studies for testing LNT. Aside from the fact that case-control studies do not have the statistical power to test LNT in the low dose region, this ignores the inherent weaknesses in treatments of CF in case-control studies. An individual's risk of lung cancer depends on a multitude of factors on a molecular, cellular, intercellular, hormonal, etc. level that are not understood, not readily measurable, and therefore not considered in these studies. There are also a large number of potential CFs that could be, but

are not included because of time, cost, or other practical limitations. In practice, case-control studies treat only a very few CFs, often using multivariate regression which is a process of limited validity, and frequently depending on marginal statistics.

Our study has many important advantages over these case-control studies. It treats a far wider diversity of CF, and even includes a strong argument that an unidentified CF cannot be important—no such arguments are available in case-control studies which can easily be rendered invalid by an unrecognized CF. Our study largely avoids use of multivariate regression with its inherent weaknesses pointed out in Sec. 2.2 above and elsewhere (Kleinbaum, Kupper, and Muller 1988). It includes a method for treating cases where no data are available on a required variable, by use of “plausibility of correlation”. It includes a wide variety of geographic areas and population characteristics, whereas case-control studies are normally confined to a single, or at most a few local areas. Statistical uncertainties, one of the greatest limitations in many case-control studies, are virtually eliminated.

It should be understood that the success in treating confounding factors reported here is due to a combination of fortunate circumstances not present in the great majority of studies. The very large number of data points, 1601 counties, with good quality data for each on hundreds of different variables, is highly unusual. But perhaps more important is the fact that radon levels in homes are very weakly correlated with most other variables like climate, socioeconomics, ethnicity, etc that might affect lung cancer rates.

One might question the fact that our study leans heavily on plausibility arguments, especially plausibility of correlation. But case-control studies choose the few CFs they investigate and the control groups they adopt based solely on subjective judgments of plausibility. The CF analyzed in our study include all of the many hundreds that are available.

With all these advantages, the problem in accepting our study is difficult to understand unless someone can suggest a plausible specific CF that could possibly explain our discrepancy. It must be specific in order to address the issue of plausibility, but it is not necessary to show that it does explain our discrepancy, only that it might possibly explain it. If such a suggestion is forthcoming, I would be eager to address it. I have tried very hard to solicit such a suggestion but have had only marginal success. Therefore, until such a CF is suggested, it seems reasonable to conclude that LNT fails our experimental test and must therefore be invalid in the low dose region covered by Fig. 1.

## REFERENCES

- Cohen, B.L. Variation of radon levels in U.S. homes correlated with house characteristics, location, and socioeconomic factors. *Health Phys.*60:631–642;1991.

B. L. Cohen

- Cohen, B.L. Test of the linear-no threshold theory of radiation carcinogenesis for inhaled radon decay products, *Health Phys.* 68:157–174;1995.
- Cohen, B.L. Response to Lubin's proposed explanations of our discrepancy. *Health Phys.* 75:18–22;1998a.
- Cohen, B.L. Response to criticisms of Smith et al. *Health Phys.* 75:23–28;1998b.
- Cohen, B.L. Updates and extensions to tests of the linear-no threshold theory, *Technology* 7:657–672; 2000a.
- Cohen, B.L. Testing a BEIR-VI suggestion for explaining the lung cancer vs radon relationship for U.S. Counties. *Health Phys.* 78:522–527;2000b.
- Darby, S., Deo, H., and Doll, R. A parallel analysis of individual and ecological data on residential radon and lung cancer in south-west England. *J. R. Statist. Soc.* A164, Part 1, 193–203;2001.
- Greenland, S. and Robins, J. Ecologic studies—biases, misconceptions, and counterexamples. *Am. J. Epidemiol.* 139:747–760;1994.
- Kleinbaum, D.G., Kupper, L.L., and Muller, K.E. "Applied Regression Analysis and Other Multivariable Methods". PWS-Kent Publishing Co. (Boston),1988.
- Lagarde, F. and Pershagen, G. Parallel analyses of individual and ecologic data on residential radon, cofactors, and lung cancer in Sweden. *Am J Epidemiol* 149:268–274;1999.
- Lubin, J. and Boice, J. Lung cancer risk from residential radon; meta-analysis of eight epidemiologic studies. *J Nat'l Cancer Inst* 89:49–57;1997.
- Lubin, J.H. On the discrepancy between epidemiologic studies in individuals of lung cancer and residential radon and Cohen's ecologic regression. *Health Phys.* 75:4–10;1998.
- Morgenstern, H. Ecologic studies in epidemiology: concepts, principles, and methods. *Annual Rev. Public Health* 16:61–81;1995.
- NAS (National Academy of Sciences) Committee on Biological Effects of Ionizing Radiation. *Health risks of radon and other internally deposited alpha emitters (BEIR-IV)*. Washington, DC, National Academy Press, 1988.
- NAS (National Academy of Sciences) Committee on Biological Effects of Ionizing Radiation. *Health effects of exposure to radon (BEIR-VI)*, Washington, DC, National Academy Press, 1999.
- Stidley, C.A. and Samet, J.M. Assessment of ecologic regression in the study of lung cancer and indoor radon. *Am. J. Radiol.* 139:312–322;1994.