Paper

TEST OF THE LINEAR-NO THRESHOLD THEORY OF RADIATION CARCINOGENESIS FOR INHALED RADON DECAY PRODUCTS

Bernard L. Cohen*

Abstract—**Data on lung cancer mortality rates vs. average radon concentration in homes for 1,601 U.S. counties are used to test the linear-no threshold theory. The widely recognized problems with ecological studies, as applied to this work, are addressed extensively. With or without corrections for variations in smoking prevalence, there is a strong tendency for lung cancer rates to decrease with increasing radon exposure, in sharp contrast to the increase expected from the theory. The discrepancy in slope is about 20 standard deviations. It is shown that uncertainties in lung cancer rates, radon exposures, and smoking prevalence are not important and that confounding by 54 socioeconomic factors, by geography, and by altitude and climate can explain only a small fraction of the discrepancy. Effects of known radon-smoking prevalence correlations—rural people have higher radon levels and smoke less than urban people, and smokers are exposed to less radon than non-smokers—are calculated and found to be trivial. In spite of extensive efforts, no potential explanation for the discrepancy other than failure of the linear-no threshold theory for carcinogenesis from inhaled radon decay products could be found.**
Health Phys. 68(2):157–174; 1995

Key words: radon; carcinogenesis; lungs, human; radiation, low-level

INTRODUCTION

The cancer risk from low level radiation is normally estimated by use of a linear-no threshold theory (with or without added terms that apply at higher doses). This theory is a logical consequence of the widely accepted view that a single particle of radiation interacting with a single cell nucleus can initiate a cancer; the number of initiating events is then obviously proportional to the number of particles of radiation, and hence to the dose. However, there is nothing in this line of reasoning about the role of biological defense mechanisms that prevent the billions of potential initiating events we all experience from each developing into a fatal cancer. If exposure to low level radiation were to stimulate these biological defense mechanisms, that effect would be added to the effect of linear-no threshold, and could cause a radical deviation of observed effects from the predictions of that theory alone in the low dose region.

There is now a substantial body of evidence indicating that low level radiation does indeed stimulate such biological defense mechanisms (Luckey 1991; Sugahara et al. 1992; Calabrese 1994). For example, it has been shown (Shadley and Wolfe 1987) that human lymphocyte cells previously exposed to low level radiation suffer fewer chromatid breaks when later exposed to large radiation doses, and this effect has been traced to stimulated production of repair enzymes by the low level radiation (Wolfe 1992). Similar effects have been demonstrated *in vivo* for bone marrow cells and spermatocytes in mice (Cai and Liu 1990). In addition to reducing chromosome aberrations, pre-exposure to low level radiation has been found to reduce induction of mutations (Sanderson and Morely 1986; Kelsey et al. 1991) and to increase survival rates (Shadley and Dai 1992; Azzam et al. 1992) in cells later exposed to high radiation doses. Low dose pre-exposure of drosophila reduces the number of dominant lethal mutations induced by later high dose radiation (Fritz-Niggli and Schaeppi-Buechi 1991). Low dose radiation has also been shown to stimulate immune functions in mice as measured by PFC (plaque-forming cell) reaction, MLC reaction (mixed lymphocyte culture, used as a test for T-cell function), reaction to Con A (concanavalin-A, a lectin that stimulates T-lymphocytes), NK (natural killer cells, which recognize and kill tumor cells) activity, and ADCC activity (anti-body dependent cell mediated cytotoxicity, which assists NK activity) in splenocytes (Liu 1992).

All of this evidence surely leaves linear-no threshold open to serious question in its applications to low level radiation. These applications have had tremendous societal consequences—adding over 100 billion dollars to the cost of U.S. nuclear power plants and largely denying the nation the great potential benefits of that technology, leading to expenditure of a projected 150 billion dollars for clean-up of government facilities, etc.—all this in spite of the fact that linear-no

* University of Pittsburgh, Pittsburgh, PA 15260.
(*Manuscript received* 29 *July* 1993; *revised manuscript received* 10 *June* 1994, *accepted* 1 *August* 1994)
0017-9078/95/$3.00/0

threshold theory has never been verified with experimental data in the low dose, low dose rate region of all the important applications. Clearly, then, it is of utmost importance to further seek such verification. That is the purpose of this paper.

LUNG CANCER RATES VS. INDOOR RADON LEVELS

A compilation has recently been completed of average indoor radon levels in 1,729 U.S. counties, over half of all U.S. counties and representing nearly 90% of the total U.S. population (Cohen 1992, 1994). Data from it were used to derive Fig. 1 (a and b) which show plots of age-adjusted lung cancer mortality rates, *m*, for white males (1a) and females (1b) (Riggan and Mason 1983) vs. average radon level, *r*, in living areas of homes in these counties. Radon levels are given in the widely used units of $r_0 = 37$ Bq m^{-3} (1.0 pCi L^{-1}). Rather than showing a data point for each county, all counties within various ranges of *r* (marked on the base line) are grouped together, and only the average *m* for each group is plotted, along with the standard deviation of the mean (error bars) and the first and third quartiles. The solid line is the best straight line fit to the data for the individual counties. In Fig. 1 (a and b), we see a clear tendency for *m* to decrease with increasing *r*, in sharp contrast to the increase expected from the fact that radon is believed to cause lung cancer.

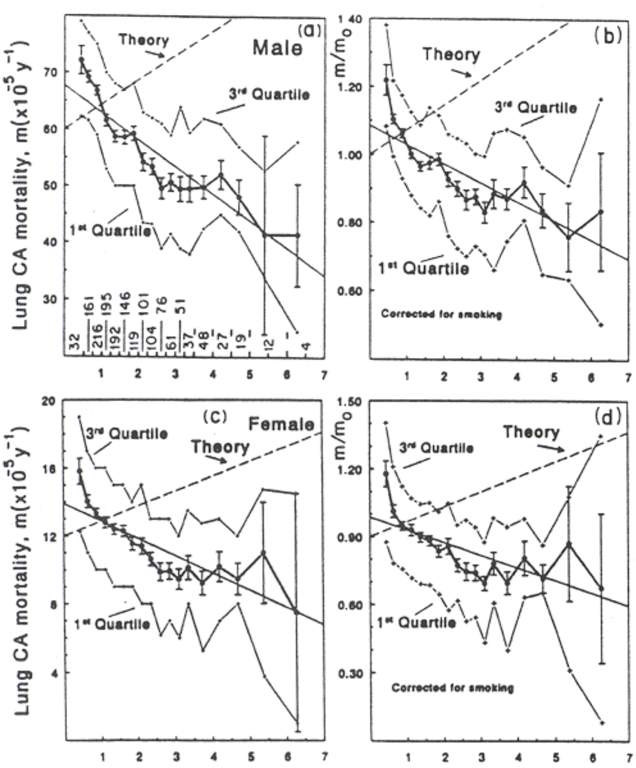

Fig. 1. Lung cancer mortality rates vs. mean radon level for 1,601 U.S. counties. Data points shown are average of ordinates for all counties within the range of *r*-values shown on the baseline of Fig. 1a; the number of counties within that range is also shown there. Error bars are standard deviation of the mean, and the first and third quartiles of the distributions are also shown. Fig. 1c, d are lung cancer rates corrected for smoking prevalence (m/m_0) vs. radon level [eqns (5) and (6)]. Theory lines are arbitrarily normalized lines increasing at a rate of 7.3%/r_0.

These data, although they are potentially explainable in many ways, are the starting point of a test of the linear-no threshold theory. A preliminary report of this test (based on data for 965 counties from a single data source) has been published previously (Cohen and Colditz 1994). There have been about 50 studies of the relationship between radon exposure and lung cancer reviewed recently by Neuberger (1991, 1992). Of these, 13 have involved measurements of radon levels including 7 ecological studies, 4 case-control studies, and 2 that involve both. A later report on one of the studies has appeared recently (Schoenberg 1992) and at least one ecological study was not included in the review (Haynes 1988). A recent review (Stidley and Samet 1993) discusses problems with these ecological studies. There have been two recent case-control studies (Perslagen 1994; Letourneau 1994) that have attracted substantial attention but neither of these gives useful information in the region below 5_{r_0} where essentially all data in Fig. 1 are contained.

A. The ecological fallacy

Fig. 1 is an example of what epidemiologists call an "ecological study," which means that it compares the average mortality rates for groups of people (i.e., populations of counties), *m*, with their average exposure, *r*. This is quite different from what epidemiologists normally study, the mortality risk to an individual, *m'*, vs. that individual's exposure to radon, *r'*. To illustrate the difference, let us suppose that only radon levels above 20 r_0 can cause cancer, and that county X has $r = 1.5\ r_0$ but no exposures above 20 r_0 while county Y has $r = 1.0\ r_0$ but 1% of its population is exposed above 20 r_0; then county Y would have the higher cancer rate, *m*, even though county X has the higher average radon exposure, *r*. To state the problem succinctly, the average exposure does not determine the average risk. This is the major contributor to what is called "the ecological fallacy" (Robinson 1950; Selvin 1958).

However, it has been shown (Cohen 1990a, 1990b) that the above problem does not apply when there is a linear-no threshold relationship. This is familiar to health physicists from the well known fact that "person-rem," an ecological quantity, determines the number of deaths in a cohort, regardless of how the dose is distributed among the individuals in the cohort. Expressing this loosely in terms of our problem, the cohort is the population of a county; dividing the number of deaths by the population gives the mortality rate, *m*, and dividing person-rem by the population gives the average exposure, *r*. This indicates crudely why m vs. r data like those shown in Fig. 1 are not affected by the above problem. A more complete demonstration is given in the next section, and it is refined through the rest of the paper.

Other problems arising from the fact that this is an ecological study, often called "ecological bias," will be discussed in section P.

B. Smoking prevalence, S

An obvious potential explanation for the unexpected pattern of Fig. 1 (a and b) is that there may be a strong tendency for the prevalence of cigarette smoking, S, to be higher in low radon counties than in high radon counties; that is, there is a large negative r-S correlation.

The effect of smoking prevalence is most conveniently treated by use of the BEIR IV theory (NAS 1988), which gives separate risk estimates for smokers and non-smokers. First let us consider non-smokers only, for which we use the subscript n. According to BEIR IV, the mortality risk for individual i, m_i, of age A_i, living in a house with radon level r_i, in a given year is

$$m_i' = a_n g(A_i)[1 + h(A_i) r_i'] \quad (1)$$

where g and h are given in BEIR IV as a function of age, and a_n is a constant inserted here for convenience in normalizing.

We next sum both sides of eqn (1) over i and divide by the number of individuals in the cohort, q. The sum over m_i' gives the total number of deaths among non-smokers, and dividing this by q gives the mortality rate from lung cancer among non-smokers, m_n. The sum over the right side of eqn (1) is evaluated in Appendix A. Dividing that result by q and equating it to m_n then gives

$$m_n = a_n (1 + br) \quad (2)$$

where $b = 10.8\%/r_0$.

We now turn to consider smokers. Let p = the number of smokers in the county and we repeat the above process for smokers to obtain analogously

$$m_s = a_s (1 + br).$$

Note that we assume here that the average radon level, r, is the same in houses of smokers and non-smokers; this assumption is relaxed in section L. Note also that, as discussed in Appendix A, the constant b is the same for smokers and non-smokers; this assumption is peculiar to the BEIR IV model and it is relaxed in section M.

The total lung cancer mortality rate for the county, including the number of deaths among smokers, $p\, m_s$, and among non- smokers, $q\, m_n$, with their sum divided by the total population of the county, $(p + q)$, is

$$m = (pm_s + qm_n)/(p + q).$$

In terms of the smoking prevalence, $S = p/(p+q)$, this reduces to

$$m = [Sa_s + (1 - S)a_n]\,(1 + br) \quad (3)$$

C. Migration

One obvious problem in our test is that people move frequently and therefore do not spend their whole lives—and receive all of their radon exposure—in their county of residence at time of death, where their death is recorded and contributes to mortality rates. This problem has been treated in some detail previously (Cohen 1990b, 1992a, 1993); the procedure selected was to assume that people spend a fraction of their lives, f, in their county of residence at death, and the remaining fraction, $(1 - f)$, in areas of U.S. average radon level, $\bar{r}$. With the exception of Florida (FL), California (CA), and Arizona (AZ), where many people move to retire, all areas of the U.S. have $f > 0.5$, and the national average is $f = 0.7$ (Cohen 1992a). The three retirement states (FL, CA, AZ) were deleted from the study, reducing the total number of counties to 1,601. (This has no significant effect on the results of the analyses that follow.) Eqn (3) is modified by the assumption to

$$m = [Sa_s + (1 - S)a_n](1 + 0.3b\bar{r} + 0.7br) \quad (4)$$

The migration problem is thereby handled to a reasonable approximation by modifying the theory, and no correction to the data is necessary. Dividing both sides of eqn (4) by $[Sa_s + (1-S)a_n]\,(1+0.3\bar{r})$ and inserting the BEIR IV values of the constants (Cohen and Colditz 1994) converts this to

$$m/m_0 = 1 + Br \quad (5)$$

where $B = 7.3$

$$m_0 = (9 + 99S) \quad \text{males}$$
$$m_0 = (3.7 + 32S) \quad \text{females,} \quad (5a)$$

where B is in units of $\%/r_0$ [%/37 Bq m^{-3} (%/pCi L^{-1})] and m_0 is in units of deaths y^{-1} per 100,000 population. In order to compare data with theory, one should, therefore, plot m/m_0 vs. r. Such plots are shown in Fig. 1 (c and d); the methodology for estimating S to determine m_0 for each county is given in Cohen and Colditz (1994) and will be discussed further below (section G). Fig. 1 (c and d) may be viewed as Fig. 1 (a and b) corrected for variations in smoking prevalence. Note that correcting for smoking does little to explain the unexpected behavior of the data.

Further discussion considers plots of m/m_0 vs. r (like those in Fig. 1c and 1d), concentrating on the slope, B; BEIR IV theory predicts $B = +7.3$ (eqn 5a). For comparison with observations, the best fit of the data to

$$m/m_0 = A + Br \quad (6)$$

is used to derive values of A and B. One might use B/A in eqn (6) as the quantity to be compared with B in eqn (5), but this would be equivalent to normalizing the line through the data to the theory line at $r = 0$. However, it would be more appropriate to normalize these lines in the region where most of the data points lie, at about $1.2\,r_0$ (Fig. 1), which is crudely equivalent to using $B/0.9A$. To simplify the problem, B-values from eqns (6) and (5) are compared directly, ignoring the small differences between $0.9A$ and 1.00 (typically < 10%).

The best fits to the data points for the 1,601 counties in Fig. 1 (c and d) give $B = -7.3 \pm 0.6$ for males (± is one standard error) and $B = -8.3 \pm 0.8$ for females, discrepant by about 20 standard deviations with the values expected from BEIR IV theory, $B = +7.3$. This difference will be called "the discrepancy," and the remainder of this paper deals with the attempts to explain it. If data from Florida, California, and Arizona were not deleted, results would be $B = -6.5 \pm 0.7$ for males and $B = -9.0 \pm 0.8$ for females.

It is immediately clear that the discrepancy cannot be explained by statistics; the probability for chance discrepancies of 20 standard errors is truly negligible.

D. Uncertainties in radon data

The compilation of radon data used here derives from three independent sources:

- University of Pittsburgh measurements (PITT);
- Measurements by U.S. Environmental Protection Agency (EPA); and
- Data bases compiled by individual states, not based on PITT or EPA (STATE).

The data used in Fig. 1, TOTAL, is an average of as many of these three as are available for each county. The evidence for the reliability of these data has been discussed extensively (Cohen 1992c,1994b).

The results for B from the various data sets treated separately are

DataSet	*Counties*	*B(male)*	*B(female)*
PITT	1,151	−6.4	−9.1
EPA	1,074	−6.4	−6.3
STATE	358	−6.8	−10.8

as compared with $B = -7.3$ for males and $B = -8.3$ for females from TOTAL. For the 663 counties included in both PITT and EPA, the results for B are

PITT	663	−5.8	−7.3
EPA	663	−5.5	−6.7

For the 296 counties in both PITT and STATE, the results for B are

PITT	296	−6.4	−13.4
STATE	296	−6.3	−12.1

Note that each of the independent data sets gives essentially the same result, especially when an identical selection of counties is used. This gives a high degree of confidence that errors in the values of r are not responsible for the discrepancy, or for any significant fraction of it.

By studying correlations between values of r for specific counties from the separate data sources, estimates of uncertainties in the individual r-values were derived (Cohen 1992c). They are generally about 17% (one standard error). If these uncertainties are random in direction—and it is difficult to imagine reasons why they might be otherwise—this should bias our values of B toward the null by about 17% (Fuller 1987). For example, the B value for males should be corrected for this bias from −7.3 to −8.5, further increasing the discrepancy. This correction will not be used in the following discussion, but it would easily compensate for most of the small corrections to be considered later that might reduce the negative value of B.

One problem common to PITT, EPA, and STATE is that people who live in apartment buildings, where radon levels are generally low, are under-represented. To investigate this problem, data are available on H5, the percentage of housing units that are in buildings containing 5 or more units (U.S. Census 1982). An extreme correction for this problem would be to assume that none of these are included in the radon surveys and that all of them have zero radon concentrations. This correction changes B from −7.3 to −7.1 for males, and from −8.3 to −8.4 for females.

Another approach to this problem is to delete the counties with large H5. If the 20% with the largest H5 were deleted (H5 > 16.6%), B would be changed to −7.2 for males and −7.9 for females. If the 40% with the largest H5 were deleted (H5 > 8.9%), B-values would become −6.7 and −6.7, respectively. If the 60% with the largest H5 were deleted (H5 > 5.7%), B-values would be −6.2 and −7.3, respectively. It thus seems clear that the "apartment problem" is not an important cause of the discrepancy.

The slope of a regression line can often be heavily and unduly influenced by the effects of a few outlying data points. To investigate this effect, various indices suggested in the statistics literature were used for discarding 10 and 20 outliers—leverages (L), studentized residuals (S), Cook's distance (C), DFITS (D), and residuals (R) in the regression of m/m_0 vs. r. The results for the slope, B, are listed in Table 1. In all cases, discarding outliers increases the negative value of B, making the discrepancy worse. Outliers were not discarded in further analyses.

Table 1. Effects of discarding outliers. These are values of B obtained if 10 or 20 counties are discarded on the basis of various indices: L (leverages), S (studentized residuals), C (Cook's distance), D (DFITS), and R (residuals).

	Discard 10		Discard 20	
Index	Male	Female	Male	Female
L	−7.8	−8.7	−7.7	−9.3
S	−7.7	−8.9	−7.6	−9.3
C	−7.3	−9.1	−7.7	−9.1
D	−7.8	−9.4	−8.3	−10.0
R	−7.8	−8.9	−7.7	−9.3

An unrelated problem with the accuracy of the r-values derives from the number of measurements on which they are based. It is shown at the end of section H that this does not affect the results.

E. Sampling issues

One might wonder whether the discrepancy can be explained by peculiarities in the sample of U.S. counties under study. One way to test this is to break the data into subsets and determine B independently for each subset. Table 2 shows the results for 10 randomly selected subsets each of 800, 400, and 200 counties. The results are always reasonably close to those for the entire data set, $B = -7.3$ for males and -8.3 for females. Even a study of 200 counties would clearly show the discrepancy. In that sense, this work might be considered as eight separate studies, each leading to the same conclusion.

Dividing the data into subsets on bases other than random selection will be discussed later in this paper but again the discrepancy is invariably encountered.

One might wonder how unexpected it is to find such a large and statistically robust correlation between m and r as we find for lung cancer even if there is no causal explanation for it. In a separate project, regressions of m on r, and of m on r and S, for 33 different cancer sites were studied (Cohen 1993). Whether or not S is included, the number of standard deviations by which B differs from zero, and the coefficient of determination, R^2, were found to be at least 2.7 times larger for lung cancer than for any other cancer type, and for all but two types they were at least 4 times larger. One must therefore conclude that the strong m-r correlation seen in Fig. 1 is a truly unusual and remarkable occurrence, and therefore should not be dismissed as something that might occur by chance with reasonable probability.

F. Uncertainties in lung cancer mortality rates, *M*

Lung cancer mortality rates (Riggan and Mason 1983) are derived from mortality records for 1970–1979 which are the latest age-adjusted rates available at this time [Cohen and Colditz (1994) presents a crude analysis using more recent data; it makes the discrepancy larger]. No attempt was made to analyze uncertainties arising from variations in the efficiency of collecting these data, but it is difficult to imagine reasons why these variations might correlate with radon levels other than through geography as a confounder, a topic treated later in this paper.

One problem that can be treated is that arising from statistical uncertainties; low population counties had relatively few lung cancer deaths in 1970–1979. The distributions of lung cancer mortality rates, m, for all counties within a narrow range of radon levels have typical relative standard deviations of 24% for males and 34% for females. If the statistical uncertainty for m is no more than half this large, its statistical accuracy may be judged to be irrelevant; this requires at least 69 deaths for males and 35 deaths for females, which are roughly the numbers expected in counties with populations of 23,000 and 58,000, respectively.

As a test of this problem, all counties with at least these populations were given equal weight, while counties with lower populations were given weights inversely proportional to the variance of m, which is just proportional to their population. With this criterion, weights are reduced for 465 of the 1,601 counties for males, and for 994 counties for females. When this weighting was used in the regression of m/m_0 vs. r to determine the slope B, the values of B were changed from -7.3 ± 0.6 to -7.1 ± 0.5 for males and from -8.3 ± 0.8 to -7.4 ± 0.7 for females. Since these changes are relatively small, this weighting was not used in our other studies.

Table 2. Values of B derived from 10 randomly selected subsets of 800, 400, and 200 counties. Bottom row is the standard deviation of the mean, determined from the 10 values listed.

	Males			Females		
Select No.	800	400	200	800	400	200
1	−6.9	−6.5	−5.8	−6.7	−5.4	−4.8
2	−7.6	−7.8	−7.8	−7.8	−6.8	−6.5
3	−7.8	−6.9	−7.0	−7.5	−8.9	−10.7
4	−7.2	−7.9	−8.3	−7.4	−10.5	−9.2
5	−7.1	−6.4	−5.6	−10.1	−10.9	−10.0
6	−7.8	−8.0	−7.0	−6.9	−5.4	−4.2
7	−6.2	−5.8	−5.5	−8.1	−9.0	−8.0
8	−6.9	−7.8	−5.0	−8.4	−6.4	−6.6
9	−7.3	−6.4	−7.1	−10.1	−10.3	−12.7
10	−8.7	−10.9	−8.5	−9.9	−9.6	−11.9
Average	−7.4	−7.4	−6.8	−8.3	−8.3	−8.5
SD	0.2	0.4	0.4	0.4	0.7	0.9

G. Uncertainties in smoking prevalence, *S*

Direct information on smoking prevalence is available only by state, with the best data derived from a 1985 survey by Bureau of Census (U.S. PHS 1990). This was corrected to the appropriate time period for deaths occurring in 1970–1979 by use of data on the time variation of the national smoking prevalence (U.S. PHS 1987) assuming that the relative prevalence for each state remained unchanged. This gives the smoking prevalence for each state, S'. It was then assumed that the S-value for a county was S' times a correction factor for the fraction of the county population that lives in an urban area, PU; this correction factor was derived from a regression of lung cancer rates on PU and was found to be remarkably constant for all regions of the nation.

This procedure gives a distribution of S-values within a single state about half as wide as the distribution of S-values for the various states. To investigate suggestions that this may seriously under-represent the variations of S for counties within a state, the above correction factor was doubled in magnitude; this changed B from −7.3 to −7.0 for males and from −8.3 to −8.0 for females, a trivial effect.

There are three obvious problems in this method of deriving S-values:

(a) The direct data are derived from a 1985 survey, which is inappropriate for predicting lung cancer mortality in 1970–1979.
(b) The direct data are on smoking prevalence in states rather than in counties.
(c) There is no consideration of intensity of smoking, degree of inhalation, use of filters, etc.

Problem (a) and part of problem (c) can be avoided if we derive S' values from state cigarette sales tax collections, which are available on a yearly basis (Tobacco Institute 1988). If these are taken to be proportional to S' for males, and the method outlined above is used to derive S-values, the results are

- 1975: $B = -8.3 \pm 0.7$;
- 1970: $B = -9.0 \pm 0.6$; and
- 1960: $B = -10.1 \pm 0.7$.

The discrepancy is larger for this source of data than for that used in Fig. 1 ($B = -7.3$). This source was not used in the other studies.

An alternative approach for deriving S-values that avoids all three of the above problems is to use lung cancer rates in counties. This, of course, must be done in a way that is independent of r, which is accomplished by stratifying the data on the basis of radon levels, r, into six subsets. All counties in a given subset therefore have approximately the same r-value.

Since variations in S are principally due to socioeconomic factors, values of S for each county are estimated from a linear combination of socioeconomic variables (SEV). The 54 SEV available for each county that are used here and in later parts of our analysis are listed in Appendix B; they are basically all of those in County and City Data Book—1988 (U.S. Census 1988) that are not intrinsically proportional to population, plus a few others, including population and population density. The original S' is also included as an SEV. A scoring system was then developed to determine which of these SEVs are most useful for predicting m for each of the six subsets of counties. Out of a possible 384 points, the highest scores for males were S': 369; HA: 295; PU: 164; EW: 123; GR: 107; and SC: 99 (no others were > 69). The highest scores for females were S': 319; EF: 227; SC: 146; GR: 125; PT: 118; EW: 103; and PU: 101 (no others were > 94).

Multiple regressions of m on these SEV were done for each of the six subsets, and coefficients for each SEV were recorded. In all but one case (GR for females, which was therefore not used), these coefficients were reasonably consistent among the subsets, and average coefficients were derived. These average coefficients were then used to determine S-values for each county from values of its SEV. Note that these S-values do not suffer from any of the problems, (a), (b), (c), listed above. When these S-values are used in eqn (5a) to determine m_0 for use in eqn (6), negative B-values are reduced from -7.3 to -6.0 for males and from $B = -8.3$ to -6.3 for females. They thus give only a minor reduction in our discrepancy.

Since using m-values to determine a parameter, S, to be used in fitting m-values is a somewhat questionable procedure, the original S-values were used for further studies. However, this exercise gives confidence that problems (a), (b), (c) listed above are not responsible for much of the discrepancy.

Nevertheless, the values of S being used are subject to substantial uncertainty, leaving open the possibility that errors in S can somehow explain the discrepancy. This would be the case if the true S-values had a much stronger negative r-S correlation than those being used. To quantify this potential effect, the S-values for the 1,601 counties were reassigned in perfect reverse order of their r-values. This "perfect" negative r-S correlation gives a coefficient of correlation (CORR) between r and S of -0.96 for males and -0.92 for females. When these reassigned S-values are used in eqn (5a) to calculate m_0 for use in eqn (6), the results are $B = +0.7$ for males and $B = -0.3$ for females. The negative slopes are eliminated but only about half of the discrepancy is explained; these B-values are far short of $+7.3$ expected from theory.

While this perfect negative r-S correlation is a drastic assumption, one can go even further and broaden the distribution of S-values for the 1,601 counties. The characteristics of this distribution for males (S in percent) are as follows: mean = 51.7; standard deviation (SD) = 6.9; and min/max = 25.5/69.8. As a broadened distribution, the S-value for each county is taken to be twice as different from the mean. This gives a distribution with mean = 51.7; SD = 13.8; and min/max = 0/88. These S-values are then reassigned to counties in perfect reverse order of their r-values to obtain a "perfect" negative r-S correlation—S(perfect)—as before: alternatively these S-values are reassigned to counties randomly to obtain S (random). S is then taken to be

$$S = G\, S(\text{perfect}) + (1 - G)\, S(\text{random}),$$

where G is a parameter that can be varied to obtain any desired negative coefficient of correlation (CORR-r) between S and r. To eliminate the negative slope in Fig. 1 (i.e. to make $B = 0$) is found to require CORR-$r = -0.64$, and to obtain the theory value, $B = +7.3$, requires CORR-$r = -0.90$. Recall that these results are based on the drastic assumption that the width of the distribution of S-values is twice as large as in the best estimates. An analysis which is independent of the width of the distribution of S-values is presented at the end of section M.

If one is completely skeptical about the methods used here to estimate S-values, an alternative approach is to assume that the distribution of S-values is the same as the distribution of lung cancer mortality rates, m, for males (aside from a normalizing factor to give the correct national average for S), and calculate the CORR-r required to explain the discrepancy. This would seem to give an upper limit on the width of the S-distribution since other factors must contribute to the width of the m-distribution. Utilizing the methods of the previous paragraph—combining S(perfect) and S(random) in various ratios—it is found that obtaining the theory value, $B = +7.3$, requires CORR-$r = -0.91$, and just eliminating the negative slope to make $B = 0$ requires CORR-$r = -0.62$.

Consideration was next given to the likelihood of such strong r-S correlations. Since there is no apparent direct causal relationship between r and S, the most likely source of r-S correlations is confounding by socioeconomic variables, SEV. It therefore seems reasonable to expect the r-S correlation to be similar to the correlation of r with these SEV.

The largest magnitude correlation with r (|CORR-r|) for any of the 54 SEV is 0.37 for EF which is clearly an urban vs. rural effect which will be treated in detail in section L, and this effect may also explain all of the five largest |CORR-r|. The second largest |CORR-r| is 0.30, and for 49 of the 54 SEV it is less than 0.23. For the S-values being used here, CORR-$r = -0.28$ for males and -0.19 for females; for the S-values derived from cigarette sales tax data, CORR-$r = -0.16$ for 1975, -0.16 for 1970, and -0.11 for 1960. For the S-values derived from lung cancer rates, CORR-$r = -0.40$ for males and -0.34 for females.

From these examples, it seems clear that the negative r-S correlations cannot be nearly as strong as those needed to reduce B to zero, let alone to produce the large positive B predicted by the theory. It is therefore reasonable to conclude that uncertainties in S are not a very important cause of our discrepancy.

H. Confounding factors that correlate with socioeconomic variables (SEV)

If the theory is correct, the only reasonable explanation for the discrepancy is that there are one or more confounding factors that correlate strongly and with opposite signs with both m and r. They thereby introduce a strong negative correlation between m and r which is not due to a direct causal relationship. Smoking was the obvious candidate because of its known strong correlation with lung cancer, but it was considered in great detail above, and found not to explain the discrepancy. The next most obvious type of confounder would be socioeconomic variables (SEV). Consideration was therefore given to each of the 54 SEV listed in Appendix B as a possible confounding factor (CF).

If a particular SEV is an important CF, stratifying the data on it into subsets would greatly reduce the problem as each subset (i.e., each stratum) would have approximately the same value of the CF. The average of the B-values obtained from analysis of each of the various subsets would then give a value of B free from the effects of confounding.

The data are stratified into quintiles of 1,601/5 = 320 counties each, on the basis of each of the 54 SEV in turn. This gives 54 × 5 (quintiles) × 2 (sexes) = 540 subsets, each analyzed to derive a value of B. The results are shown in Table 3. Note there that all 540 B-values are negative. Thus, the negative slopes in Fig. 1 (c and d) are found if we consider only the most urban counties, or if we consider only the completely rural counties; if we consider only the richest counties or if we consider only the poorest; if we consider only those with the best medical care, or if we consider only those with the worst medical care; if we consider only the most rapidly growing counties or only the counties with declining populations; and so forth for each of the 54 SEV. These negative slopes are also found for all the strata in between, as for example, if we consider only counties with close to the national average income, or close to average education, or medical care, or any one of the other 51 SEV.

Averages over the five quintiles for each SEV and sex are shown in the last four columns of Table 3 along with their t-ratios, the number of standard deviations by which B differs from zero. The average value of B for the five quintiles, which is a determination of B free of confounding by that SEV, vary for males from -5.6 to -7.7 with a mean of -6.9 ± 0.5, and for females from -5.4 to -9.1 with a mean of -7.7 ± 0.8, quite close to the values for the entire data set, -7.3 and -8.3, respectively.

Thus, it is clear that no one of the SEV is an important enough CF to explain more than a tiny fraction of the discrepancy. Since SEV are normally strongly correlated with some other SEV, this probably means that no SEV is an important enough CF to substantially reduce the discrepancy. In fact, no factor which correlates strongly with any of the SEV can explain much of the discrepancy as that SEV would act as a surrogate for it in analyses like those done here. For example, air pollution cannot be an important CF since it correlates strongly with several of the SEV.

The last two rows of Table 3 show the results of stratifying on the number of radon measurements in the PITT and EPA data bases. The consistency of these results indicates that insufficient numbers of radon measurements in some counties is not an important source of difficulty.

I. Confounding by combinations of socioeconomic factors

Another possibility is that some combination of SEV may act cooperatively to confound the M-r relationship (where $M = m/m_0$). The best available approach to investigating this question is through multiple regression analysis, taking the relationship for each county to be

$$M = m/m_0 = A + Br + C_1X_1 + \ldots + C_{54}X_{54}, \quad (7)$$

where $X_1, X_2, \ldots, X_{54}$ are the values of the 54 SEV and A, B, $C_1, C_2, \ldots, C_{54}$ are the constants chosen to

Table 3. Values of *B* obtained from stratifying the data into quintiles ($Q_1, \ldots, Q_5$) on the basis of each of the SEV. Note from the column headings that minus signs and decimal points have been deleted to save space. The last four columns are averages of values for the five quintiles.

	−10B for males					−10B for females					Average male		Average female	
SEV	Q-1	Q-2	Q-3	Q-4	Q-5	Q-1	Q-2	Q-3	Q-4	Q-5	*B*	*t*	*B*	*t*
PT	63	74	65	44	55	105	54	44	32	87	−6.0	−4.9	−6.4	−3.8
PD	59	67	64	73	28	65	89	64	71	58	−5.8	−4.72	−6.9	−4.05
PI	51	69	46	85	116	97	102	65	74	39	−7.3	−6.04	−7.5	−4.29
PU	71	63	76	72	69	90	67	83	69	92	−7.0	−5.87	−8.0	−4.88
PW	74	54	55	47	49	73	18	66	35	83	−5.6	−3.67	−5.5	−2.65
PS	50	62	31	77	82	71	73	46	80	54	−6.0	−4.33	−6.5	−3.42
PE	94	47	51	72	92	57	21	73	85	136	−7.1	−5.89	−7.4	−4.26
PO	105	50	60	12	93	85	1	71	33	121	−6.4	−4.54	−6.2	−3.29
PY	81	71	46	38	132	109	100	115	44	24	−7.4	−6.05	−7.8	−4.84
PN	68	92	57	74	71	111	93	74	75	47	−7.2	−6.00	−8.0	−4.57
PH	67	81	57	45	127	93	104	102	24	76	−7.5	−6.25	−7.9	−4.69
VB	41	55	54	82	124	81	66	91	79	97	−7.1	−5.97	−8.3	−4.72
VC	85	50	25	68	66	106	100	40	65	67	−5.9	−4.71	−7.6	−4.04
VD	73	85	65	54	86	65	50	76	77	133	−7.3	−6.16	−8.0	−4.60
VI	56	66	48	88	96	77	93	49	70	98	−7.1	−5.83	−7.7	−4.33
VM	80	44	76	51	84	70	100	75	88	54	−6.7	−5.48	−7.7	−4.38
VS	74	78	49	64	57	102	101	76	11	28	−6.4	−5.29	−6.4	−3.52
VP	65	69	81	72	68	84	99	55	44	89	−7.1	−5.89	−7.4	−4.28
VH	81	79	59	51	93	38	100	76	71	136	−7.3	−6.05	−8.4	−5.05
SS	95	60	67	51	89	65	47	70	76	130	−7.2	−6.02	−7.8	−4.43
SC	83	40	62	72	81	69	97	39	74	81	−6.8	−5.65	−7.2	−4.05
SH	73	91	61	40	77	83	82	76	109	107	−6.8	−7.56	−9.1	−6.47
SU	64	50	98	69	84	69	75	80	96	99	−7.3	−6.09	−8.4	−4.73
SE	39	44	80	74	103	58	67	76	103	118	−6.8	−5.65	−8.4	−4.85
HO	78	85	46	51	113	91	79	58	81	86	−7.5	−6.07	−7.9	−4.53
HA	74	54	59	51	71	65	56	53	83	77	−6.2	−5.02	−6.7	−3.75
HV	98	49	50	73	90	47	83	107	85	99	−7.2	−6.01	−8.4	−4.87
HN	64	48	97	80	92	37	107	80	102	79	−7.6	−6.13	−8.1	−4.55
EI	79	88	66	61	70	73	91	95	58	82	−7.3	−6.13	−8.0	−4.56
EH	92	84	69	47	80	71	114	89	74	62	−7.4	−6.17	−8.2	−4.68
EJ	45	38	77	96	102	77	88	87	141	43	−7.2	−5.97	−8.7	−4.98
EV	45	44	65	105	102	86	92	69	143	52	−7.2	−5.96	−8.8	−5.08
EU	97	64	47	59	89	115	76	82	76	27	−7.1	−5.83	−7.5	−4.46
EW	59	81	59	55	76	95	63	72	62	42	−6.6	−5.39	−6.7	−3.76
EP	89	76	75	58	85	37	112	56	67	164	−7.7	−6.28	−8.7	−5.01
EM	78	87	76	66	14	91	148	129	36	13	−6.4	−5.34	−7.8	−4.67
ER	48	82	59	83	93	65	69	83	85	97	−7.3	−6.20	−8.0	−4.65
ES	74	64	68	73	79	64	71	80	86	97	−7.2	−6.09	−8.0	−4.76
EG	54	91	70	66	81	88	62	60	89	104	−7.2	−6.10	−8.1	−4.58
EF	75	53	70	55	48	64	83	34	13	75	−6.0	−4.77	−5.4	−3.01
EA	62	48	51	78	67	108	35	65	109	63	−6.1	−5.04	−7.6	−4.27
EL	68	92	80	61	43	100	144	92	61	29	−6.9	−5.64	−8.5	−4.83
ED	58	51	72	87	84	64	67	81	95	81	−7.0	−5.93	−7.8	−4.52
EC	77	72	61	67	82	91	84	75	44	106	−7.2	−6.19	−8.0	−4.74
EX	72	61	68	80	81	77	96	92	51	89	−7.2	−6.22	−8.1	−4.71
GF	74	32	94	80	81	55	28	70	132	107	−7.2	−6.04	−7.8	−4.44
GL	58	43	64	65	116	53	62	91	54	137	−6.8	−5.78	−7.9	−4.50
GE	86	73	55	68	91	96	118	61	85	52	−7.5	−6.23	−8.2	−4.75
GH	113	65	61	66	55	38	101	106	104	64	−7.2	−5.97	−8.3	−4.90
GP	77	47	67	105	83	93	92	67	19	88	−7.5	−6.13	−7.2	−4.14
GW	134	71	70	30	70	75	117	54	42	115	−7.5	−5.77	−8.1	−4.37
GR	61	81	64	64	65	28	47	37	112	99	−6.7	−5.47	−6.5	−3.62
GJ	68	54	64	62	91	77	83	42	107	92	−6.8	−5.63	−8.0	−4.59
GV	86	71	73	61	62	68	90	78	93	83	−7.1	−5.92	−8.2	−4.65
NP	77	54	60	80	60	52	100	122	114	76	−6.6	−5.01	−9.3	−4.77
NE	56	69	59	87	86	55	88	60	36	85	−7.1	−4.93	−6.5	−3.21

obtain the best fit to the data for all counties. One might worry about utilizing 56 adjustable constants, but with 1,601 pieces of data, the fitting procedure gives values with small statistical uncertainties, including $B = -3.1 \pm 0.6$ for males and $B = -3.5 \pm 0.9$ for females. These reduce the discrepancy with theory by 29% and 31%, respectively. One might therefore conclude that this is an effect of some importance.

However, the statistics literature contains frequent warnings about use of many variables in a multiple regression to quantify the causal relationship of one particular variable. An obvious problem here is that, since there is a strong negative correlation between M and r as indicated by Fig. 1, any SEV that has a strong correlation with M is likely to have a strong correlation of opposite sign with r. In fitting eqn (7), its term will therefore "drain away" some of the strength of the Br term, reducing the value of B. With many such variables acting in that way, the value of B may be substantially reduced.

This problem was investigated in some depth utilizing the data for males, beginning with a determination of the coefficients of correlation (R) of each SEV with M (CORR-M) and with r (CORR-r). These are plotted in Fig. 2, where a clear pattern is evident—there is a very strong tendency for an SEV with a large CORR-M to have a large CORR-r of opposite sign.

To study the problem further, the number of SEV is first truncated to reduce computational labor by keeping only the 13 SEV in Fig. 2 that have $|\text{CORR-}M| > 0.125$ (points encircled). It is these SEV which we expect to be most important in the problem outlined above. In fact, if only these 13 SEV are retained in eqn (7), B is changed from −7.3 (simple regression) to −3.7 ± 0.6, which is 86% of the reduction from keeping all 54 SEV.

A model is then introduced based on artificial SEV which have a built-in CORR-M but no built-in CORR-r. The SEV are first "standardized"—let X = [X-mean(X)]/[SD(X)]—and several hundred artificial SEV, SEV(art), are generated for each county as

$$\text{SEV(art)} = p\,M + (1 - |p|)\ \text{sample-}M, \qquad (8)$$

where sample-M is a random sample of the M-values from the 1,601 counties (each used only once) and p is given several different values between −0.25 and +0.25. Five sets of 13 SEV(art) are then selected from this list such that in each set, there is one SEV(art) which has the same CORR-M as one of the actual 13 SEV. Note that these SEV(art) have no built-in CORR-r; any CORR-r they may have must derive only from the correlation between M and r.

When these sets of SEV(art) are used in eqn (2), the values of B obtained are −5.2 ± 0.5, −5.1 ± 0.5, −4.9 ± 0.5, −5.0 ± 0.5, and −4.9 ± 0.5. Thus, utilizing these SEV(art) which have no built-in CORR-r, reduces B from −7.3 to −5.0, which is 64% of the reduction obtained from using our actual SEV. An alternative procedure based on 17 SEV by adding points surrounded by squares in Fig. 2, selected because of their large CORR-r, gave a reduction of B from use of these SEV(art) that is 73% of the reduction obtained by use of the actual SEV. It therefore seems reasonable to conclude that about two-thirds of the 29% reduction in our discrepancy obtained from use of multiple regression is due to the methodological problem under investigation. Thus the actual effect of confounding by combinations of SEV is to reduce the discrepancy by perhaps 10%.

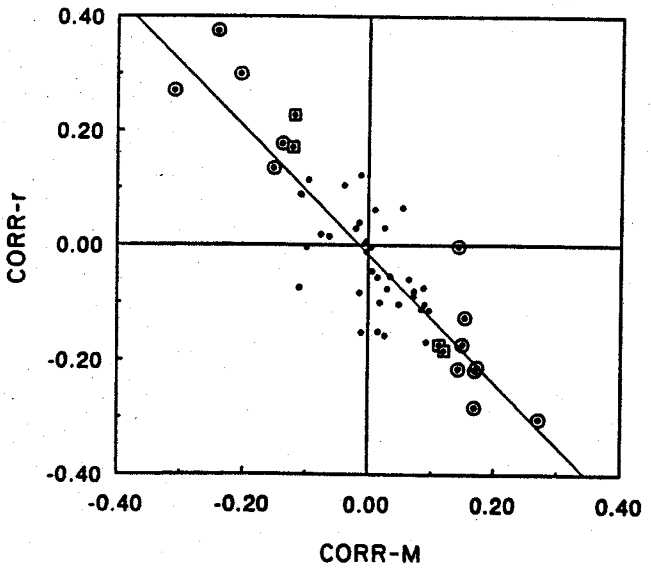

Fig. 2. CORR-r vs. CORR-M for the 54 socioeconomic variables (SEV). Circles and squares around some of the points are to refer to those points in the text.

J. Confounding by geography

The only thing known to correlate strongly with r is geography (Cohen 1991), which suggests that it be considered as a CF. It is therefore treated by the stratification method used above for SEV. The U.S. Bureau of Census divides the nation into four regions, each consisting of two or three divisions. In the previous study (Cohen and Colditz 1994) it was found that stratifying on divisions reduced the discrepancy for males and females by 22% and 16%, and stratifying to the level of individual states reduced it by 28% and 17%, respectively.

Table 4. Results of stratifying on geography by regions and by divisions. The bottom two rows are the averages of results for the four regions and for the nine divisions.

Region Division	Number of counties	Male B	Male t	Female B	Female t
Northeast	215	−6.0	−5.72	−9.9	−5.63
New England	65	−3.0	−0.81	−0.5	−0.06
Mid Atlantic	150	−6.1	−5.11	−11.6	−6.01
North Central	612	−5.1	−6.10	−5.4	−4.80
East NC	308	1.2	1.02	1.6	1.06
West NC	304	−6.0	−4.81	−5.8	−3.26
South	566	−7.9	−6.02	−8.7	−4.71
South Atlantic	273	−9.2	−3.90	−5.2	−1.58
East S. Central	135	−1.9	−1.07	−6.8	−2.54
West S. Central	158	−14.4	−6.02	−15.3	−4.47
West	208	−5.5	−3.36	−8.1	−2.86
Mountain	167	−1.3	−.069	−6.7	−1.83
Pacific[a]	41	0.6	.150	−9.3	−1.43
Averages					
Regions		−6.1	−5.30	−8.0	−4.50
Divisions		−4.4	−2.21	−6.6	−2.23

[a] Includes only WA and OR.

In the present analysis, there are data for nearly twice as many counties, which greatly improves the statistics. The results of stratifying on regions and on divisions are listed in Table 4. Averaging over B values for the 4 regions reduces B from −7.3 to −6.1 for males, and from −8.0 to −6.6 for females, reducing the discrepancy by 8.2% and 9.0% respectively. Averaging B-values for the 9 divisions reduces B from −7.3 to −4.4 for males, and from −8.3 to −6.6 for females, reducing the discrepancy by 19.9% and 10.9% respectively, somewhat less than in the previous study.

In the present data base, there are 34 states with at least 20 counties having known radon levels, as compared with only 18 states in the previous study. The results for individual states are listed in Table 5. Averaging over B-values for these 34 states gives B = −6.1 for males and −7.2 for females, which now reduces the discrepancy by only 8.2% and 7.1% respectively, far less than in the previous study.

Table 5. Results of stratifying on geography to the level of individual states. All states with at least 20 counties in the data file are included.

State	Number	Male B	Male t	Female B	Female t
AL	38	−5.3	−1.05	−1.5	−0.16
AR	36	−14.6	−2.96	−13.8	−1.85
CA	44	−2.9	−0.47	−19.6	−1.71
CO	41	−1.4	−0.42	−8.9	−1.84
FL	33	−1.2	−0.23	3.5	0.65
GA	52	−12.4	−2.46	−5.9	−0.65
ID	40	2.1	0.45	−8.2	−1.04
IL	54	−3.1	−1.07	−5.2	−1.31
IN	58	0.0	0.01	1.7	0.54
IA	98	−3.8	−1.99	−5.0	−1.52
KS	21	−12.1	−1.85	−16.1	−2.20
KY	32	−3.2	−1.07	−10.7	−2.63
LA	21	−20.8	−0.73	−42.6	−0.90
MD	24	−11.5	−1.83	−12.2	−2.14
MI	53	3.2	1.02	0.3	0.09
MN	64	−2.1	−1.40	−10.1	−3.84
MO	31	−9.6	−1.19	−6.8	−0.57
NB	43	−2.3	−0.50	−2.7	−0.48
NJ	21	−0.3	−0.09	−2.6	−0.49
NM	30	−2.9	−0.60	−2.1	−0.19
NY	62	0.7	0.25	−8.2	−1.46
NC	54	−9.1	−2.54	2.8	0.64
ND	39	−5.3	−1.41	22.6[a]	2.60
OH	88	0.0	−0.02	−0.2	−0.10
OK	43	−12.6	−1.87	−24.1	−2.60
PA	67	−1.5	−1.07	−5.3	−2.32
SC	36	−37.9	−2.43	−0.7	−0.03
TN	46	0.4	0.16	−4.8	−1.18
TX	58	−10.4	−3.14	−11.7	−2.56
VA	66	−6.1	−1.43	−15.0	−1.93
WA	29	−5.4	−1.55	−12.8	−1.85
WV	37	0.2	0.05	−10.4	−1.91
WI	55	−9.6	−2.05	−4.9	−0.70
WY	21	−5.9	−1.02	−4.9	−0.37
Averages		−6.1	−1.08	−7.2	−1.07

[a] This is dominated by a single rural county in which there was one death over the 10-y period. If that county is deleted, $B = 5.5, t = .074$.

Perhaps the most important point is that in the previous study it seemed like the finer the stratification on geography, the less the discrepancy became. But now with a much larger data base giving much better statistics, that trend is reversed, with the finest stratification reducing the discrepancy by only about 8%.

Since our S-values for counties are calculated as the S-value for the state plus a correction for "%-urban", it has been suggested that our analyses for individual states does not adequately adjust for smoking variations among the counties of the state. Since the correction is derived from lung cancer rates, there is good reason to believe that it is reliable. However, as a test of sensitivity to this problem, the correction was doubled as in the related test discussed in section G. The effects on Table 5 were relatively minor; the average of the B-values (bottom line of Table 5) was changed from −6.1 to −5.7 for males, and from −7.2 to −6.4 for females.

K. Confounding by physical features

Apart from smoking, socioeconomics, and geography, another type of potential confounding factor is physical features of the location like barometric pressure (which is determined by altitude), temperature, precipitation, etc. Unfortunately, no data on these were found for counties, but there are data for the most important cities in each state (U.S. Census 1982), which are averaged to obtain values for states. To use these, the project must be reconsidered as dealing with states rather than with counties.

To be consistent with the study of counties, data for Arizona, California, and Florida are deleted, and because of a misunderstanding, Alaska and Hawaii were also deleted, but in retrospect this may be justified by the fact that, in many ways, they are not typical of the other states. That leaves 46 states (including DC), a very small number of data points in comparison with the 1,601 in the study of counties, which means that statistics are a potential problem.

Applying the treatment used for counties to this data set gives $B = -13.0 \pm 2.3$ for males and -14.4 ± 2.7 for females. These are much larger negative values than were obtained from the data on counties, making the discrepancy larger, but they agree well with the B-values for states in Cohen and Colditz (1994), −12.0 and −14.4, derived from a much smaller data base on radon levels.

The physical features considered here are altitude above sea level, average winter temperature, average summer temperature, centimeters of precipitation per year, number of days per year with more than 0.25 mm precipitation, average wind speed, and percent of time with sunshine as compared with the maximum possible. The data are stratified into three equal groups of 15–16 states, those with the lowest, highest, and mid-range values of the feature under consideration, and the data in each group are fit to eqn (6) to determine B.

Table 6. Results from stratifying data for states on the basis of physical features, listed in column (1). See explanation in text.

				Males			Females		
Stratify on . . .	Ranks	Min	Max	*B*	SD	R^2 (%)	*B*	SD	R^2 (%)
Altitude (m)	1–15	5.5	65	−6.1	4.8	4	−21.9	9.5	30
	16–31	65	140	−7.1	3.9	19	−8.5	5.3	16
	32–46	158	632	−13.7	3.4	55	−16.3	4.5	50
	1–31	5.5	140	−8.9	3.4	18	−13.0	4.2	25
	39–46	320	632	−16.0	7.6	42	−8.1	6.5	9
Temp.-Jan. (°C)	1–15	−13.2	−3.0	−11.1	3.7	41	−13.9	4.1	47
	16–31	−2.3	0.8	−18.4	4.2	58	−19.5	4.1	61
	32–46	1.4	11.6	−14.5	5.4	35	−1.9	6.3	1
	1–7	−13.2	−6.7	−9.8	4.9	44	−7.3	3.2	51
	40–46	5.2	11.6	−45	16	61	−64	17	74
Temp.-July (°C)	1–15	19.5	22.7	−17.3	4.4	54	−18.2	5.5	46
	16–31	22.8	25.1	−14.9	3.0	63	−18.9	2.8	77
	32–46	25.6	46.2	−14.4	5.3	36	−3.7	1.8	2
Precipitation (cm y^{-1})	1–15	18.3	77.7	−11.7	3.7	44	−15.3	5.3	39
	16–32	78.2	103.6	−5.1	3.6	11	−10.7	4.5	28
	33–46	108.0	170.2	−12.2	6.8	20	−20.4	8.5	33
	1–8	18.3	41.1	−10.5	6.2	33	−11.9	11.9	15
Precipitation (d y^{-1} >0.25 mm)	1–16	51	105	−10.5	3.4	41	−12.8	4.3	39
	17–31	110	124	−9.0	6.9	12	−22.0	7.9	37
	32–46	124	154	−14.7	6.4	29	−11.7	7.8	15
	1–8	51	90	−13.6	2.3	85	−7.4	8.1	12
	9–16	95	105	−12.4	5.6	44	−14.5	5.4	56
Wind (m/s average)	1–15	2.9	3.9	−13.6	7.2	21	−13.7	8.6	17
	16–30	4.0	4.4	−12.5	3.9	43	−9.8	5.4	20
	31–46	4.5	5.8	−13.6	3.6	50	−17.1	3.6	63
Sunshine (%)	1–15	49	57	−14.7	5.1	40	−16.4	7.1	29
	16–31	57	62	−12.1	3.4	48	−17.5	3.4	66
	32–46	62	81	−12.7	4.2	41	−5.2	5.7	6

Results are shown in Table 6. For example, the first section involves stratifying on altitude and the top row gives results for the 15 states with the lowest altitude (ranks 1–15), in which the minimum and maximum altitudes are 5.5 m and 65 m above sea level; the data for males in this group is fit by $B = -6.1 \pm 4.8$ (one standard error, SE), and R^2, the percent of the variation in the data explainable by the simple linear relationship, is 4%; the data for females is fit by $B = -21.9 \pm 9.5$, with $R^2 = 30\%$. In addition to the stratification into three equal groups with lowest, mid-range, and highest values of the feature under consideration which appear in the top three rows of each section, results for other groupings are also shown in some of the sections where these are judged to be useful. For example, in the fourth section of Table 6 on precipitation, the values for the 15 states with lowest precipitation ranged from 18.3 to 77.7 cm y^{-1}, a very wide variation; a grouping of the 8 states with the lowest precipitation, ranging from 18.3 to 41.1 cm y^{-1} is therefore added.

A total of 28 groupings is included in Table 6. For both sexes, this gives 56 values of *B* and all 56 are negative. The average value of *B* for the three equal size groups for males [females] is −9.0 [−15.6] for altitude, −14.7 [−11.8] for winter temperature, −15.5 [−13.6] for summer temperature, −9.7 [−15.5] for mm y^{-1} of precipitation, −11.4 [−15.5] for d y^{-1} of precipitation, −13.2 [−13.2] for wind speed, and −13.2 [−13.0] for %-sunshine. In no case do these deviate from the values without stratification, −13.0 [−14.4], in the same direction for both males and females, and in no case is the average deviation for the two sexes more than 0.6 standard deviations.

These studies therefore lead to the conclusion that none of the physical features is an important confounding factor in the relationship between lung cancer and radon exposure. The strong decrease in lung cancer rates corrected for smoking frequency with increasing radon exposure is found if we consider only the low altitude states or if we consider only the high altitude states; if we consider only the warmest or if we consider only the coldest; if we consider only the wettest or only the driest; etc. It is also found if we consider only states where these features are close to average.

L. Effects of recognized *r-S* correlations

Extensive studies have been previously reported on how house characteristics, locations, socioeconomic factors, etc., correlate with radon levels in homes (Cohen 1991). There were two of these factors that one would expect to correlate with lung cancer incidence: (1) urban houses have 25% lower radon levels than rural houses, and urban people smoke more frequently; and (2) houses of smokers have 10% lower radon levels than houses of non-smokers, which is a direct radon-smoking correlation on the level of individuals (as opposed to a correlation on the level of counties). Other potential confounders could be considered here, such as income or education: poor and less educated people smoke more than average. However, little difference was found in mean radon levels as a function of income, value of house, or education, and the relationships are not monotonic—e.g., middle income people have the highest radon levels. Furthermore, these factors should be taken into account by the numerous SEV.

To calculate the effects of factors (1) and (2), a model is introduced in which it is assumed that the BEIR IV formula gives the correct lung cancer risk, and that variations in smoking prevalence are determined only by urban vs. rural considerations. While this model is highly oversimplified, it includes all the elements relevant to the effects we are studying. It only makes use of the data on average radon levels, r, and the percentage of population living in urban areas, PU, for each county.

It is assumed that the mean radon levels are xr_m for rural areas and r_m/x for urban areas of a county, where r_m is the measured value for the county as a whole. For males, it is assumed that the smoking frequency, S, is 0.5 y in urban areas (where 0.5 was the national average smoking frequency) and 0.5 y^{-1} in rural areas. (For females 0.5 is replaced by 0.32.) Both x and y are normally greater than unity, and they are treated as parameters. For x, the best estimate is 1.12. Regression analysis on data for counties indicates that as PU goes from 0–100%, male lung cancer rates go from 53–66 y^{-1} per 100,000, a variation of ± 10% from the average; this is interpreted to indicate proportional prevalences of smoking, or y = 1.10.

In order to treat factor (2), it is assumed that the above values of r are multiplied by z for smokers and divided by z for non-smokers. The best estimate is z = 1.05.

The derivation of eqn (3) treated two groups, smokers and non-smokers, but here this is increased to four groups, urban smokers (US), urban non-smokers (UN), rural smokers (RS), and rural non-smokers (RN). The percentage, P, in each category is

$$P(\mathrm{US}) = \mathrm{PU}(0.5\,y)$$
$$P(\mathrm{UN}) = \mathrm{PU}(1\text{–}0.5\,y)$$
$$P(\mathrm{RS}) = (1\text{-PU})\,(0.5\,y^{-1})$$
$$P(\mathrm{RN}) = (1\text{-PU})\,(1\text{–}0.5\,y^{-1}).$$

The average radon level for each category is

$$r(\mathrm{US}) = r_m/xz$$
$$r(\mathrm{UN}) = r_m z/x$$
$$r(\mathrm{RS}) = r_m x/z$$
$$r(\mathrm{RN}) = r_m xz.$$

The average radon level for the county, to be used in the regression, is then $r = P(\mathrm{US})r(\mathrm{US}) + P(\mathrm{UN})r(\mathrm{UN}) + P(\mathrm{RS})r(\mathrm{RS}) + P(\mathrm{RN})r(\mathrm{RN})$, and the smoking prevalence, S, is $S = P(\mathrm{US}) + P(\mathrm{RS})$. We can then calculate m_0 from eqn (5).

The mortality rate for each category is

$$m''(\mathrm{US}) = a_s\,[1 + .073r(\mathrm{US})]$$
$$m''(\mathrm{RS}) = a_s\,[1 + .073r(\mathrm{RS})]$$
$$m''(\mathrm{UN}) = a_n\,[1 + .073r(\mathrm{UN})]$$
$$m''(\mathrm{RN}) = a_n\,[1+.073r(\mathrm{RN})],$$

and the county mortality rate to be used in the multiple regression is $m = P(\mathrm{US})m''(\mathrm{US}) + P(\mathrm{UN})m''(\mathrm{UN}) + P(\mathrm{RS})m''(\mathrm{RS}) + P(\mathrm{RN})m''(\mathrm{RN})$.

Note that the radon level r is not quite the same as the measured value, r_m, but for x, y, z not very different from unity, it is close. The only important thing for our model is that the distribution of r-values for all counties is realistic.

Once m, r, and m_0 have been calculated for each county, the slope, B, of the regression of m/m_0 on r and the slope, B', of the regression of m on r can be determined. The values of B and B' for males and females are listed in Table 7. In that Table, section A, $x = y = z = 1$, is the baseline situation, section B gives the urban-rural effect, section C gives the effect of the radon-smoking correlation, and section D combines both of these effects. The top entry in each section is based on the best estimates (BE), and succeeding lines treat deviations from unity of 2 × BE, 4 × BE, 8 × BE, and 16 × BE.

We see that the urban-rural effect reduces B' by about 18% with the BE, and for more than 4 × BE it reduces the slope to zero, but the correction for smoking, using m/m_0, compensates for these effects causing B to be almost unaffected.

Factor (2), the radon-smoking correlation on the level of individuals, has a much lesser effect on B', reducing it only to 60% of its baseline positive value even if the effect is 16 × BE. But it has a stronger effect on B since it is not related to smoking prevalence and hence is not compensated by our smoking correction. The combination of the two effects studied here gives roughly what might be expected from a linearly independent relationship.

Table 7. Results for model that tests effects caused by known r-S correlations. See discussion in text.

Section	x	y	z	B (male)	B (female)	B' (male)	B' (female)
A	1	1	1	7.0	7.0	4.1	.98
B	1.12	1.1	1	7.0	7.0	3.3	.81
	1.24	1.2	1	6.9	6.9	2.2	.59
	1.48	1.4	1	6.8	6.7	0.23	.02
	1.96	1.8	1	6.5	6.4	−2.5	−.41
C	1	1	1.05	6.7	6.7	3.9	.94
	1	1	1.1	6.5	6.4	3.8	.89
	1	1	1.2	6.0	5.8	3.5	.82
	1	1	1.4	5.1	5.0	3.0	.70
	1	1	1.8	3.9	3.9	2.3	.54
D	1.12	1.1	1.05	6.7	6.7	3.1	.77
	1.24	1.2	1.1	6.3	6.3	1.9	.51
	1.48	1.4	1.2	5.6	5.6	−0.3	.07
	1.96	1.8	1.4	4.5	4.7	−2.8	−.40
	2.92	2.6	1.8	3.4	3.9	−3.6	−.56

It is important to recognize here that the BEs are based on a great deal of data and hence are reasonably accurate, and anything greater than 2 × BE is highly unlikely. This means that B is probably reduced by only 5% as a result of these effects, and a reduction by more than 10% is highly unlikely. The BEIR IV prediction is thus reduced only from +7.3 to about +6.9, which contributes very little to explaining the large negative values of B obtained from the actual data.

Perhaps the most important aspect of this section is that the effects calculated here are typical of the largest effects that can be reasonably expected from confounding factors. Their very small impact on the discrepancy leads to an impression that no confounding relationship can be reasonably expected to resolve that discrepancy.

M. Linear-no threshold theories other than BEIR IV

Up to this point we have considered only the BEIR IV theory, but other linear-no threshold theories have been proposed based on the miner data and differing principally in their treatment of smoking. How specific are the discrepancies reported here to the details of the BEIR IV theory?

A more general form of eqn (1) is

$$m' = a + br' \qquad (9)$$

where both a and b may have different values for smokers and non-smokers. If we proceed as in deriving eqn (3) from eqn (1), but ignoring the correction for migration which is the same for all theories, this leads to

$$m = Sa_s + (1 - S)a_n + [Sb_s + (1 - S)b_n]r$$

which can be re-written

$$m = Sa_s + (1 - S)a_n + [eS/\bar{S} + (1 - e)]b'r \qquad (10)$$

where $\bar{S}$ is the national average value of S, and e, b′ are new constants replacing b_s, b_n. By setting the expressions for m in eqns (10) and (5) equal to each other, we find that eqn (5) is a special case of eqn (10) with b′ = 4.9, e = 0.85 for males, and b′ = 1.17, e = 0.73 for females. The parameter e is an index of the relative risk from radon to smokers and non-smokers; for e = 0 their risks are equal and for e = 1, only smokers are at risk. The true values of e within the range 0 to 1 are a matter of substantial uncertainty, and this fact is clearly acknowledged in the BEIR IV Report. An important advantage of eqn (10) is that the national average value of m is independent of e, as we can see by setting $S = \bar{S}$.

The first two terms in eqn (10) represent risks of smoking unassociated with radon exposure, a matter that has been thoroughly studied and is subject to little uncertainty. The factor b′ in eqn (10) is derived directly from the miner data, with little sensitivity to the value of e, and it is therefore not uncertain by more than about 50%. Thus, by varying e between 0 and 1, and varying b′ over a relatively small range, we should include any linear-no threshold theory based on the miner data. We proceed by using the data on males and determining the ratio of observed m to values of m calculated from eqn (10), o/c, for each state. We then determine the slope, B'', of the best straight line fit to o/c vs. r. If theory is correct, it should be zero.

When we set b′ equal to the BEIR IV value and decrease e in steps from 1 to 0, B'' varies only from −0.16 to −0.17. There is essentially no sensitivity to the value of e.

We then set e equal to the BEIR IV value, 0.84, and multiply b′ by a factor f. As f is decreased from 2 to 1 to 0.5 to zero, the slope B'' changes only from −0.195 to −0.168 to −0.148 to −0.120. If we restrict our consideration to the maximum expected variation of ±50%, the variation is only between −0.184 and −0.148.

We conclude that our discrepancy between observation and theory would apply with only minor differences to any linear-no threshold theory derived from the miner data. The reason for this is easy to understand: the principal difference between various models is that they give widely different treatments of smoking, but since the r-S correlation is relatively small, this has little impact on the relationship in Fig. 1 between radon exposure and lung cancer rates.

A more direct way to avoid dependence of this study on the specific treatment of smoking in BEIR IV is to stratify the data on S and investigate each stratum as an independent data set. Since S-values are then approximately the same for all counties in the same data set, the potential for confounding by S is greatly reduced. The data are stratified into deciles, 10 sets of 160 counties each. For males, B-values in order of increasing S are −5.9, −6.7, −11.0, −5.8, −7.4, −7.6, −6.8, −6.5, −4.0, −6.0, all negative with an average of −6.8, vs. −7.3 without stratification. For females, B-values are −10.5, −4.5, −11.9, −6.7, −10.0, −4.2, −7.5, −10.6, −10.4, −6.2, again all negative and with an average of −8.2 vs. −8.3 without stratification.

Clearly, the functional form of S-dependence in the m-r-S relationship derived from the BEIR IV theory is not the cause of the discrepancy. This treatment also avoids effects of the width of the distribution of S-values. However, it does not address the problem of "intensity of smoking." A partial answer to that problem is to stratify on the S-values derived from lung cancer mortality rates described in section G. For males, this process gives an average B = −5.2 vs. −6.0 without stratification, and for females B = −5.3 vs. −6.3. These represent about an 8% reduction in the discrepancy.

N. Requirements on an unrecognized confounder

It is, of course, logically possible that there is some unrecognized confounding factor that can explain the discrepancy. This is a logical possibility in any type of epidemiological study, and few of these have included as thorough an investigation of confounders as has been done here.

However, it is interesting to consider the properties required of a confounder to resolve the discrepancy:

1. It must have a very strong correlation with lung cancer, comparable to that of cigarette smoking, but still be unrecognized.
2. It must have a very strong correlation of opposite sign with radon levels.
3. It must not be strongly correlated with any of the 54 socioeconomic variables (SEV).
4. It must be applicable in a wide variety of geographic areas and independent of altitude and climate.

How credible is the existence of such an unrecognized confounder? Requirement number 1 alone severely strains its credibility, since tremendous effort has gone into lung cancer studies. This unrecognized confounder must have increased by orders of magnitude since the beginning of this century, have been much more important in males in the first half of the century, with effects on females rapidly catching up in recent years; these are all very difficult requirements, fulfilled, to our knowledge, only by smoking. There has also been extensive study of factors that may correlate with radon levels, and other than geography, all correlations have been rather weak. There is no readily apparent reason, aside from the factors considered in section L, why any of them should correlate strongly with lung cancer rates.

Since all SEV correlate strongly with some other SEV, requirement number 3 essentially eliminates all socioeconomic variables and factors that correlate with them such as air pollution. The great majority of confounding factors that have been found to be important in epidemiology are of this type. Thus requirement number 3 is an important restriction. Requirement number 4 gives further important restrictions.

It therefore seems reasonable to conclude that the existence of an unrecognized confounding factor that would resolve the discrepancy is all but incredible. At least it is far less credible than failure of the linear-no threshold theory in the low dose-low dose rate region for radon decay products where that theory has never been verified.

P. Problems with ecological studies

In addition to the problems discussed in section A, several other potential problems with ecological studies have been pointed out by Morgenstern, Greenland, and Robins (Morgenstern 1982; Greenland and Morgenstern 1989, 1991; Greenland 1992; Greenland and Robins 1994). These authors give them names like effect modification, non-linearity and non-additivity, misclassification, divergent bias, cross-level bias, specification bias, standardization, etc. Some of these issues have been reviewed recently by Stidley and Samet (1993).

The basis for these papers is that an ecological study is not mathematically equivalent to an individual-level study, and they point out problems this can cause. For perspective, it is important to recognize that an individual-level study is certainly not mathematically equivalent to the logically correct approach, deriving a risk estimate from a complete knowledge of all causative factors and of how they interact; the problems this can cause are far more important and far less susceptible to treatment. All epidemiology studies only give what lawyers call "circumstantial evidence"—even if all evidence is absolutely correct and accurate, conclusions drawn from it are not mathematically certain. However, epidemiology is a very successful science and has saved many millions of lives. Judgements of epidemiology studies can, and must, be made on the basis of plausibility, and a highly plausible case can be built up from circumstantial evidence.

Of course, it is very important that careful consideration be given to the issues raised by the above authors. They have been examined and found not to be very important in this work. Some of these findings have been published elsewhere (Cohen 1990b, 1992b, 1994a), but they are reviewed here and in Appendix C.

The most important of these effects, called "nonlinearity," is the equivalent of the problems discussed in section A applied to confounding factors (CF)—the average value of a CF for a county does not necessarily determine its confounding effects. For example, the effects of family income as a CF may depend on those with very low income rather than on the county average income which may be influenced by a few people with high income. This problem is handled here by including as potential CF, percent below poverty level, percent unemployment, and median income, as well as average income. Other examples of this type are readily apparent from Appendix B for age, for education, etc.

Other types of examples of the problems discussed in section A applied to confounding factors are the cases treated in section M where their effects are found to be very small. Note that this problem does not apply to the principal variables in this work, r and S, because linearity with r is the theory being tested, and S arises from exact mathematics in converting risk to individuals into county mortality rates. It might also be noted that very few of the case-control studies of the radon-lung cancer relationship even consider confounding by socioeconomic variables, let alone the very wide variety of them treated here.

Other issues raised by Morgenstern, Greenland, and Robins are discussed in Appendix C. Efforts were made to develop other scenarios that might explain the discrepancy, but without success. Appeals were made to the authors of the papers cited above begging for suggestions, but none were received.

Negative slopes for m vs. r, like those observed here have also been reported from similar but much smaller studies in England (Haynes 1988) and France (Laterjet 1992; Dousset 1990).

References

Azzam, E. I.; de Toledo, S. M.; Raaphorst, G. P.; Mitchell, R. E. J. In: Abstracts of papers for the fourtieth annual meeting of the Radiation Research Society. Radiation Research 14; 1992.

Cai, L.; Liu, S. Z. Induction of cytogenic adaptive response of somatic and germ cells in vivo and in vitro by low dose X-irradiation. Int. J. Radiat. Biol. 58:187–194; 1990.

Calabrese, E. J. Biological effects of low level exposures to chemicals and radiation. Boca Raton, FL: Lewis Publishers; 1994.

Cohen, B. L. A test of the linear-no threshold theory of radiation carcinogenesis. Environ. Res 53:193–220; 1990a.

Cohen, B. L. Ecological vs case-control studies for testing a linear-no threshold dose response relationship. Int. J. Epidemiol. 19:680–684; 1990b.

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

Cohen, B. L. Percentage of lifetimes spent in area of residence at time of death. Environ. Res. 57:208–211; 1992a.

Cohen, B. L. Problems in ecological studies. Int. J. Epidemiol. 21:422–424; 1992b.

Cohen, B. L. Compilation and integration of studies of radon levels in U.S. homes by states and counties. Critical Reviews in Environmental Control 22:243–364; 1992c.

Cohen, B. L. Relationship between exposure to radon and various types of cancer. Health Phys. 65:529–531; 1993.

Cohen, B. L. In defense of ecologic studies for testing a linear- no threshold theory. Am. J. of Epidemiol. 139:765–768; 1994a.

Cohen, B. L. Indoor radon maps of the United States. Health Phys. 66:201–205; 1994b

Cohen, B. L.; Colditz, G. A. Test of the linear-no threshold theory for lung cancer induced by exposure to radon. Environ. Res. 64:65–89; 1994.

Dousset, M. Radon in dwellings. Aerobiologia 6:36–38; 1990.

Fritz-Niggli H., Schaeppi-Buechi C. Adaptive response to dominant lethality of mature and immature oocytes of D. Melanogaster to low doses of ionizing radiation: effects in repair-proficient and repair-deficient strains. Int. J. Radiat. Biol. 59:175–184; 1991.

Fuller, W. A. Measurement error models. New York: John Wiley; 1987.

Greenland, S. Divergent biases in ecological and individual level studies. Stat. Med. 11:1209–1223; 1992.

Greenland, S.; Morgenstern, H. Ecological bias, confounding, and effect modification. Int. J. Epidemiol. 18:269–274; 1989.

Greenland, S.; Morgenstern, H. Neither within-region nor cross-regional independence of exposure and covariate prevents ecological bias. Int. J. Epidemiol. 20:816–817; 1990.

Greenland, S.; Robins, J. Ecologic studies: biases, misconceptions, and counter examples. Am. J. Epidemiol. 139:747–760; 1994.

Haynes, R. M. The distribution of domestic radon concentrations and lung cancer mortality in England and Wales. Rad. Prot. Dosim. 25:93–96; 1988.

Kelsey, K. T., Memisoglu A., Frenkel D., Liber H. L. Human lymphocytes exposed to low doses of X-rays are less susceptible to radiation induced mutagenesis. Mutat. Res. 263:197–201; 1991.

Latarjet, R. Radiation carcinogenesis and radiation protection. Cancer J. 5:23–27; 1992.

Letourneau, E. G.; Krewski, D.; Choi, N.W.; Goddard, M. J.; et al. Case-control study of residential radon and lung cancer in Winnipeg, Manitoba, Canada. Am. J. Epidemiol 140:310–325; 1994.

Liu, S. Z. Multilevel mechanisms of stimulatory effect of low dose radiation on immunity. In: Sugahara, T.; Sagan, L. A.; Aoyama, T., eds. Low dose irradiation and biological defense mechanisms. Amsterdam: Excerpta Medica; 1992

Luckey, T. D. Radiation Hormesis. Boca Raton, FL: CRC Press; 1991.

Morgenstern, H. Uses of ecologic analyses in epidemiologic research. Am. J. Publ. Hlth. 72:1336–1344; 1982.

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.

Neuberger, J. S. Residential radon exposure and lung cancer: an overview of published studies. Cancer Detection and Prevention 15:435–443; 1991.

Neuberger, J. S. Errata. Cancer Detection and Prevention 16:87; 1992.

Pershagen, G.; Akerblom, G.; Axelson, O.; Clavensjo, B.; et al. Residential exposure and lung cancer in Sweden. N. Engl. J. Med. 330:159–164; 1994.

Riggan, W.; Mason, T. J. U.S. cancer mortality rates and trends:1950–1979. Washington, DC: U.S. Government Printing Office; 1983.

Robinson, W. S. Ecological correlations and the behavior of individuals. Am. Sociol. Rev. 15:351–357; 1950.

Sanderson, B. J. S.; Morely, A. A. Exposure of human lymphocytes to ionizing radiation reduces mutagenesis by subsequent ionizing radiation. Mutat. Res. 164:347–351; 1986.

Schoenberg, J. B.; Klotz, J. B.; Wilcox, H. B.; Szmaciasz, S. F. A case-control study of radon and lung cancer among New Jersey women. In: Cross, F. T., ed. Indoor radon and lung cancer: Reality or myth?, Richland, WA: Battelle Press; 1992: 889–904.

Selvin, H. C. Durkheim's suicide and problems of empirical research. Am. J. Sociol. 63:607–619; 1958.

Shadley, J. D.; Wolff, S. Very low doses of X-rays can cause lymphocytes to become less susceptible to ionizing radiation. Mutagenesis 2:95–96; 1987.

Shadley, J. D.; Dai, G. Q. Cytogenic and survival adaptive responses in G-1 phase human lymphocytes. Mutat. Res. 265:273–281; 1992.

Stidley, C. A.; Sa net, J. M.; A Review of ecological studies of lung cancer and indoor radon. Health Phys. 65:234–251; 1993.

Sugahara, T; Sagan, L. A.; Aoyama, T. Low dose irradiation and biological defense mechanisms. Amsterdam: Excerpta Medica; 1992.

Tobacco Institute. The tax burden on tobacco. Washington, DC; The Tobacco Institute: 1988.

U. S. Bureau of Census. County and city data book. Washington, DC: U.S. Government Printing Office; 1982.

U. S. Bureau of Census. County and city data book. Washington, DC: U.S. Government Printing Office; 1988.

U.S. Public Health Service. Cigarette smoking in the United States, 1986. Morbidity and Mortality Weekly Reports 36:581–584; 1987.

U.S. Public Health Service. Smoking and health: A national status report. DHHS Publication (CDC) 87-8396; 1990.
Wolff, S. Low dose exposures and induction of adaptation. In: Sugahara, T.; Sagan, L. A.; Aoyama, T. Low dose irradiation and biological defense mechanisms. Amsterdam: Excerpta Medica; 1992.

APPENDIX A
Sum Over the Right Side of Eqn. (1)

WE EVALUATE the right side of Eqn. (1), R, which is

$$R = \sum_i g(A_i)[1 + h(A_i)r_i'] \tag{A-1}$$

We divide the population into age groups, A_j, each of which includes N_j individuals all having essentially the same values of g and h in (A-1); we call these values g_j and h_j. This reduces (A-1) to

$$R = \sum_j N_j g_j + \sum_j (g_j h_j \sum_{i=1}^{N_j} r_i') \tag{A-2}$$

The average radon exposure for the group, r, is defined as

$$r = \frac{1}{N_j} \sum_{i=1}^{N_j} r_i'$$

We assume that r is independent of age, which makes it the same for all j. (The derivation could be carried out without this simplifying assumption, but it would be substantially more complex.) Thus

$$R = \sum_j N_j g_j + \sum_j r N_j g_j h_j$$

$$= (\sum_j N_j g_j)[1 + br] \tag{A-3}$$

where

$$b = (\sum_j N_j g_j h_j)/\sum_j N_j g_j \tag{A-4}$$

Note that

$$\sum_j N_j g_j = \sum_{i=1}^{q} g_i$$

which is just the total number of deaths, which is proportional to q.

We absorb the proportionality constant into a_n in Eqn (1) which is still to be evaluated. Thus, (A-3) becomes

$$R = q(1 + br) \tag{A-5}$$

In Table A-1, g_j and h_j are taken from Table 2-4 of BEIR-IV utilizing the widely used conversion factor, 1 pCi L^{-1} = 0.2WLM y^{-1} (WLM = working-level-months), which is based on 75% occupancy, and radon daughters at 50% of the equilibrium concentration with radon gas. Col. (1) lists the age ranges for the various groups, j, with A_j taken to be in the center of the range. Col. (2) is h_j from BEIR-IV expressed as the percent increase of risk per unit of radon exposure, $\%/r_0$, ($\%$/pCi L^{-1}). Col. (3) is the life table population of each age group, which is proportional to N_j. Col. (4) is the lung cancer rate for that age range which is proportional to g_j. Note that absolute values are not required here for N_j and g_j as any constant by which they might be multiplied appears in both the numerator and denominator of (A-4) and hence cancels out. Col. (5) is $N_j g_j$, Col. (3) × Col. (4); the sum of this column is 8253. Col. (6) is then $N_j g_j h_j$, Col. (5) × Col. (2) divided by 8253. In accordance with (A-4), the sum of Col. (5) gives b in units of $\%/r_0$.

In calculating a mortality rate for a county, one should use the age distribution for the population in column (3), and the result might be expected to depend on this age distribution, especially the fraction that is elderly. Using the age distribution in the U.S. population (1990) gives b = 10.82. One might think there would be important differences for differ-

Table A-1. Calculation of b from Eqn. (A-4).

(1) Age range (y)	(2) % increase per r_0	(3) Population (life table)	(4) Lung CA rate 10^{-5} y^{-1}	(5) (3) × (4)	(6) $\frac{(2) \times (5)}{8,253}$
0–15	3	1480	.0002	0	.000
15–25	7.5	980	.0009	1	.001
25–35	10.5	969	.008	8	.011
35–45	13.5	953	.125	119	.189
45–55	16.5	919	.75	689	1.37
55–65	16.3	840	2.17	1,823	3.60
65–75	7.5	687	4.10	2,817	2.56
75–85	8.5	440	4.85	2,134	2.20
85–95	9.5	150	3.68	552	.64
95–105	10.5	(30)	(3.68)	110	.14
			Sum:	8,253	10.72

ent states. The states with the highest and lowest percentage of population over age 65 (excluding Florida and Alaska, which we exclude from our data file) are Pennsylvania (14.8%) and Utah (8.2%). Utilizing the current age distribution for those states in column (3) of Table A-1 gives 10.78 for PA and 10.90 for UT. The largest difference in age distributions we might encounter is between males and females. Utilizing current national data for these gives b = 11.10 for males, b = 10.64 for females.

All of the above values of b are essentially identical well within other uncertainties in our treatment. Adopting a single value avoids a great deal of computing and makes our analysis much more transparent. We therefore adopt a single value, b = 10.8.

It is useful to recognize that Table A-1 can also be used to calculate the lifetime risk to an individual, m″. From Eqn (1), dropping the subscript i,

$$m'' = a_n \sum_{A=1}^{\infty} p(A)g(A)[1 + h(A)r'] \qquad \text{(A-6)}$$

where p (A) is the probability that the individual will be alive at age A. To evaluate this sum, we divide ages into age ranges, j, and let N'_g = number of years an individual expects to be alive in age range j, which is just p (A) times the number of years in the range.

Then (A-6) becomes

$$m'' = a_n[\sum_j N'_j g_j + r' \sum_j N'_j g_j h_j]$$

$$= a_n(\sum_j N'_j g_j)[1 + b'r']$$

where

$$b' = \frac{\Sigma_j N'_j g_j h_j}{\Sigma_j N'_j g_j} \qquad \text{(A-7)}$$

Eqn (A-7) can be evaluated with Table A-1 as was done above for Eqn (A-4), recognizing that Col. (3) is proportional to N'_j, and that the proportionality constant cancels out in Eqn (A-7). Thus, Eqn (A-4) gives the same result as Eqn (A-7), and b′ = b = 10.8.

This fact is useful because lifetime risks are tabulated in BEIR-IV, Table 2-4 (revised), and values of b′ are readily obtained there. One can see that b = b′ = 10.8 for males and females, smokers and non-smokers.

APPENDIX B
Socioeconomic Variables (SEV) Used in This Work

Population characteristics
PT = Total population
PD = Population km^{-2}
PI = % Pop. increase 1980–1986
PU = % in urban areas
PW = % white
PS = males per 100 females
PE = % age >64 y
PO = % age >74 y
PY = % 5–17 years old
PN = % born in state
PH = Persons per household

Vital and health statistics
VB = Births per 1,000 people
VC = % births to mothers <20 y
VD = Deaths per 1,000 people
VI = Infant deaths per 1,000 births
VM = Marriages per 1,000 people
VS = Divorces per 1,000 people
VP = Physicians per 100,000 people
VH = Hospital beds per 100,000 people

Social
SS = Social Sec. benefit per 1,000 people
SC = crimes per 100,000 people
SH = % high school grad.
SU = % college grad.
SE = $/cap. for education

Housing
HO = % owner occupied
HA = % with >1 automobile
HV = median value ($)
HN = % <8 y old

Economics
EI = $ per capita income
EH = Median household income ($)
EJ = % persons below poverty level
EV = % families below poverty level
EU = % unemployment
EW = average salary, wage
EP = $ per capita personal income
EM = % earnings from manufacturing
ER = % earnings from retail trade
ES = % earnings from services
EG = % earnings from government
EF = % earnings from farming
EA = average acres per farm
EL = % of manufacturing firms >100 employees
ED = $/cap. sales—food stores
EC = $/cap. sales—clothing
EX = $/cap. sales—eating, drink

Government
GF = Federal govt., $/cap.
GL = Local govt., $/cap.
GE = % local govt. expenditure—educ.
GH = % local govt. expenditure—health
GP = % local govt. expenditure—police
GW = % local govt. expenditure—welf
GR = % local govt. expenditure—roads
GJ = local govt. employment per 10,000 people
GV = % vote for lead party, 1984

NP = num. of measurements—PITT
NE = num. of measurements—EPA

APPENDIX C
Response to Issues Raised by Greenland, Morgenstern, and Robins

Greenland and Robins (1993) point out several potential problems with ecological studies. We respond to them here, referring to their examples by the numbers they use:

Non-linearity and non-additivity—examples 1–5

The non-linearity problem for confounding factors has been discussed in Sec. P, and the response to the less important non-additivity problem would be similar. There is no use of additivity here except, very obliquely, in Sec. I.

One problem here with Greenland and Robins (1993) is that, in their examples 3–5, they use a dependence of lung cancer risk on cigarettes per day smoked which is grossly different than the known dependence (Kahn 1966). They are only trying to demonstrate a mathematical inconsistency, but mathematical inconsistency is not an important issue here. The most any epidemiology study can achieve is a high degree of plausibility, and this requires use of plausible input. If one is free to concoct examples without this restriction, any epidemiological study can be shown to give arbitrarily large errors.

Measurement error—example 6

This would be a problem in our work if r were a dichotomous variable, but it is not. The effect of measurement error on r is discussed in Sec. E.

Our smoking variable, S, is a dichotomous variable, but the effects of uncertainties in it are given elaborate consideration in Sec. H.

Cross level bias—examples 7–8

These are not applicable to our work because the public knew nothing about radon in the relevant time period (prior to 1980) and therefore radon levels could not have influenced their actions.

"Misconceptions"

The non-linearity issue was discussed above in connection with examples 1–5. We have made essentially no use of the test of fit, R^2. We do not assume that using a large number of regions eliminates correlations with r—correlations with r are discussed in several sections of this paper, but especially in Sec. H, J, and M. The paper they reference (Cohen 1992b) merely pointed out that the extremely strong r-S correlations they concocted could occur by chance in the three county system they considered, but is much less likely to occur by chance in a much larger system. There is no assumption in any of our work that region is a confounder on the individual level.

Other papers by Greenland and Morgenstern referenced in Sec. P raise other issues or use different names. We discuss their application to this paper here:

Specification bias

This deals with a situation where m is not linearly dependent on r. Since we are testing the linear theory, that problem is not applicable here.

Effect modification

This deals with product terms in the expression for risk to an individual; for example, his risk, m′, might depend on (r′ × s′) where r′ and s′ are his exposure to radon and cigarettes respectively, or on (r′ × A′) where A′ is his age. They point out that such terms cannot be represented in an expression for average risk, m, by products of average values—(r × S) or (r × A) in these examples.

But there is no attempt to do that in this paper. The treatment of smoking is derived from the risk to an individual by rigorous mathematics. In averaging risks over the total population of a county, the age dependence of the risk to individuals can only lead to a dependence on the age distribution of the population, and this dependence is well handled by stratifying on the variables PE, PO, and PY listed in Appendix B.

Misclassification

This is basically another name for "measurement error" discussed above.

Standardization

The application in this paper could derive from the fact that lung cancer rates are age-standardized, while the other variables used are not. It is very difficult to image how that could be an important problem here, but in any case, the problem is removed by stratifying on our age-distribution variables, PE, PO, and PY.

Heterogeneity of exposure within regions

This cannot be a problem with a linear—no threshold theory. The distribution of exposures is completely irrelevant.

■ ■

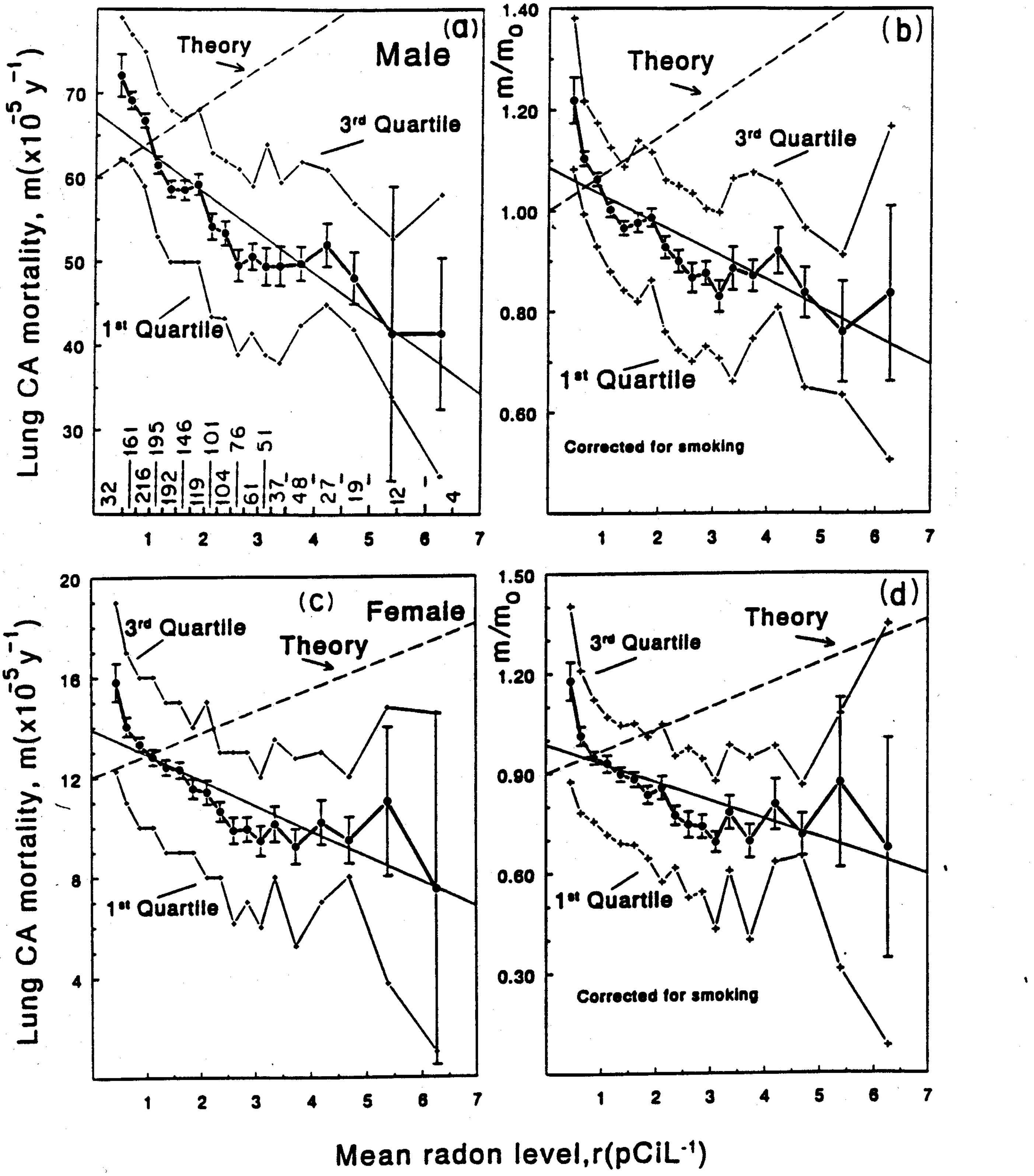
(a)
Male
Theory
3rd Quartile
1st Quartile
Lung CA mortality, m(x10^-5 y^-1)
32 216 161 192 195 119 146 104 101 61 76 37 51 48 27 19 12 4
(b)
Theory
3rd Quartile
1st Quartile
m/m_o
Corrected for smoking
(c)
Female
3rd Quartile
Theory
1st Quartile
Lung CA mortality, m(x10^-5 y^-1)
(d)
Theory
3rd Quartile
1st Quartile
m/m_o
Corrected for smoking
Mean radon level, r(pCiL^-1)

Fig 1