Low-Dose Linearity: The Rule or the Exception?

Martha Crawford* and Richard Wilson**
*Division of Applied Sciences, Harvard University, Cambridge, MA
**Department of Physics, Harvard University, Cambridge, MA

ABSTRACT

In 1976, Crump, Hoel, Langley, and Peto described how almost any dose-response relationship for carcinogens becomes linear at low doses when background cancers are taken into account. This has been used, by the U.S. Environmental Protection Agency, USEPA, as partial justification for a regulatory posture that assumes low-dose linearity, as is illustrated by a discussion of regulation of benzene as a carcinogen. The argument depends critically on the assumption that the pollutant and the background proceed by the same biological mechanism. In this paper we show that the same argument applies to noncancer end points also. We discuss the application to a number of situations: reduction in lung function and consequent increase in death rate due to (particular) air pollution; reduction in IQ and hence (in extreme cases) mental deficiency due to radiation in utero; reduction of sperm count and hence increase in male infertility due to DBCP exposure. We conclude that, although the biological basis for the health effect response is different, in each case low-dose linearity might arise from the same mathematical effect discussed by Crump et al. (1976). We then examine other situations and toxic end points where low-dose linearity might apply by the same argument. We urge that biologists and chemists should concentrate efforts on comparing the biological and pharmacokinetic processes that apply to the pollutant and the background. Finally, we discuss some public policy implications of the possibility that low dose linearity may be the rule rather than the exception for environmental exposures.

Key Words: low, dose, linearity, noncancer

INTRODUCTION

Crump et al. (1976) and Guess, Crump, and Peto (1977) reminded us that when small amount of pollutant is added to a large amount of the same pollutant, or to another pollutant operating in the same way as the first, a response linear with the
incremental dose of pollutant can result. This was a simple and elegant idea that
followed mathematically from the fact that the first derivative of a smooth curve is
finite. The idea was applied to cancer risks, and was used as a partial justification for
the U.S. Environmental Protection Agency's (USEPA) use of low-dose linearity in
regulating carcinogens. (Although the USEPA has never justified it in as much
detail as given here.) Meanwhile, the USEPA has continued to assume a nonlinear,
or "threshold," dose-response relationship for regulating chemicals that are toxic, but
not carcinogens. In this paper we demonstrate that the mathematical effect
formulated by Crump et al. (1976) applies to an array of situations in which a
pollutant causes a noncancerous, or toxic, adverse effect. The generality of the
argument suggests that a linear dose-response relationship may be the rule rather
than the exception at the low doses typical of environmental exposures, even for
ordinary toxic (noncarcinogenic) effects.

Although there may be strong biological justification for a nonlinear or threshold
dose-response relationship, the generality of the argument of Crump et al. (1976)
suggests to us that in practice, this may not have much relevance for the incremental
response to an incremental dose. If true, this would have important implications
for public policy. For example, the sheer number of situations involved suggests to the
authors that there must be some change in the way societies address these questions.
Moreover, the sharp distinction between carcinogens (which are presumed to display
linearity) and noncarcinogens (which do not), such as put forth in the Delaney
Clause of the Food and Drug Act (FDA, 1977), might well be modified by a
biological discussion of whether the pollutant at issue acts in the same way as a
background process to produce an adverse health effect.

An important difference probably remains between the low-dose risks for cancer
end points and noncancer end points. We do not know in advance, and we may
ever know, which of half a dozen heavy cigarette smokers will develop cancer
because of their habit. It is widely believed to be a purely stochastic process. This
makes it difficult to take preventive action, except for the whole population. But for
several noncancer end points examined here, it is possible, in principle, to identify
the portion of the population most at risk. This information in turn could lead to
refinements in policies and behaviors, since identification of this group would allow
a targeted approach to risk reduction. Although such an approach would likely be
less costly, especially for occupational exposures, it would raise policy and ethical
issues, which need careful discussion.

For the past two decades, regulatory agencies have sought to develop regulations
based on biological models that detail the mechanisms leading to specific effects.
However, a consideration of macroscopic patterns of dose-response relationships can
be equally important in providing windows into the black box, yielding regulations
that are effective and accurate. We make an analogy here with one of the important
subjects in physics - thermodynamics. In thermodynamics, laws of nature are
formulated macroscopically, without reference to the underlying microscopic
structure of matter, and its minute variations. Nonetheless, the laws so formulated
provide practically applicable and theoretically plausible descriptions of reality.
Indeed, for many practical applications (such as in chemical manufacturing), only the
Low-Dose Linearity

Macroscopically derived laws are actually useful, since a microscopic approach is often overly complicated and untractable (Becker, 1989). Sometimes microscopic (statistical mechanical) calculations suggest a phenomenon that is inconsistent with the macroscopic laws, and further study shows errors in the complex microscopic calculation. In this paper we are concerned with some logical practical deductions from assumed (or measured) dose–response relationships. In the distant future when biologists can correctly derive dose–response relationships from first principles, this discussion will become irrelevant, much as the thermodynamic arguments of the 19th century are being replaced by the statistical mechanical ones of the 20th century.

THE INCREMENTAL POTENCY ARGUMENT

In this section we apply the argument of Crump et al. (1976) to any medical end points, using both an analytical and a graphical approach. For the sake of argument, we assume that the biological response to a dose of a substance is inherently nonlinear, but monotonically increasing, such as depicted in Figure 1. It might for ample be represented by a power law, $R = A d^n$. ($R$ is response, $A$ is a constant, $d$ dose, $n$ is an integer greater than one.) (If the dose response is not monotonic, or $n < 1$, the situation becomes more complex and needs more careful scientific and public policy discussion. These have not yet been the focus of regulatory discussion.) Figure 1 presents the particular case of a cubic ($n = 3$). If there were no background incidence of the response in question, the dose–response relationship at low doses would be clearly nonlinear, and an infinitesimal increase of dose would give a negligible increase in response.

![Graph](image)

Figure 1. Typical nonlinear, "threshold", dose–response relationship ($R = A d^n$).
Crawford and Wilson

In general, if it is assumed that a particular background effect is caused by a mechanism similar to that caused by the pollutant at issue, it follows that there must exist an "equivalent background dose", $d_0$, to create this background response, corresponding to $R_0 = A d_0^x$ as shown in Figure 2. This "background dose" may be a background of a different substance so that it would not normally be considered, but it might be a substance that acts in the same way as the pollutant under discussion. If the dose is increased by an amount $\delta d$ due to anthropogenic (pollutant) activity, the response will increase by an amount $\delta R$, so that:

$$R_0 + \delta R = A(d_0 + \delta d)^x,$$
where $A$ is a constant greater than zero. \hfill (1)

Expanding this for $\delta d$ small (which is typical of most environmental exposures), we find:

$$R = R_0 + \delta R = A(d_0 + \delta d)^x = A d_0^x + n A (d_0^x(n-1)) \delta d + \text{Order of } (\delta d)^2 \hfill (2)$$

Solving for $\delta R$, we find:

$$\delta R = nA d_0^{x-1} \delta d$$

The incremental potency of the toxin, $f_{inc}$, is then equivalent to the slope of the dose-response curve, or:

$$f_{inc} = (nA d_0^{x-1}) \hfill (4)$$

If the dose-response curve is "anchored" by a measured acute response $R_0$, corresponding to a high dose $d_0$ (including background dose $d_0$), as is typical of data gathered through laboratory experimentation, then $R_0 = A d_0^x$.

The straight line potency, $f_{sl}$, is derived from a straight line from the response at the high anchoring dose to the origin:

$$f_{sl} = R_0 / d_0 = A d_0^{x-n}$$

whereas the incremental potency at dose $d_0$ is $f_{inc} = nA d_0^{x-1}$ \hfill (5)

The ratio of the incremental to straight line potency is then:

$$f_{inc} / f_{sl} = n \left( d_0 / d_0 \right)^{x-1} \hfill (6)$$

The incremental potency at the anchoring dose is greater than the straight line potency by the factor $n \left( d_0 / d_0 \right)^{x-1}$. In general, if the dose is close to the background dose (typical of environmental pollutants), the incremental potency will be less than the straight line potency (as determined by data at the higher dose), as can be seen from Figure 3. The value of the ratio is shown as a function of dose in Figure 4.

If the response under consideration is the probability of developing a particular cancer, the above argument is essentially the argument of Crump et al. (1976). But the argument is also applicable to noncarcinogenic responses. In application to noncarcinogens, the normal curve might describe the population variability with regard to some biological capacity, such as fertility or lung function, which compromised by the effects of a pollutant (Dockery et al., 1983; Beck, Doyle, and Schachter, 1981). We argue that, if the resultant deficiency occurs naturally via the same basic biological mechanism as that through which the pollutant exerts its
Figure 2. Threshold dose-response relationship ($R = A d^p$) with axes shifted to $R_0$ and $d_0$. Note that $\delta R_0$ is proportional to $\delta d_0$.

Figure 3. Comparison of slope derived from incremental response model, $\beta_{inc}$, and that derived by extrapolation from high-dose laboratory data, $\beta_{ch}$. Note that often $\beta_{inc} < \beta_{ch}$.
effect, then even a small increase in pollutant increases the number of people with the deficiency. Thus the pollutant elicits an incremental (and possibly linear) response with even the lowest dose, irrespective of the biological details of the dose-response relationship, provided that it monotonically increases. It should be clear, however, that nonmonotonic cases, which might exist if the distribution of the effect in the population were discrete or bimodal, need different treatment.

Often, an epidemiological measurement from an acute exposure, or a laboratory measurement from a high-dose experiment, fixes a risk ratio (RR), the ratio of the effect at the dose where it is measured $d_e$, to the background effect at the assumed background dose $d_0$. Then:

$$RR = \frac{Ad_e^n}{Ad_0^n} \text{ and } \frac{d_e}{d_0} = (RR)^{1/n}$$  \hspace{1cm} (7)

Eliminating the values of the doses from these equations, the ratio of the incremental potency to the straight line potency becomes:

$$n \left( \frac{d_e}{d_0} \right)^{(n-1)} = n (RR)^{(n-1)/n}$$  \hspace{1cm} (8)

The general requirements for this model are easily seen: (1) the origin biologically derived, dose-response curve must be smooth and monotonic; and (2) the background agent and the pollutant must act by the same mechanism or receptor in causing the effect. As a restriction upon the generality of the argument, we note that this argument involves the dose applied to the organ of concern. Any nonlinearity in the pharmacokinetic relationship between exposure and organ dose remains.

Figure 4. Incremental dose, $d$, versus Risk Ratio (RR). At low doses, when $d = d_0$, incremental potency is less than the straight line potency.
Low-Dose Linearity

It is important to realize that the coefficient of low dose linearity (the slope of the curve, starting from $R = R_0$ at $d = d_0$) will not generally be the same as the coefficient of high dose linearity, as derived from laboratory studies at acute doses $d_0$. Moreover the low-(incremental) dose slope depends not only upon the biological dose-response curve but also upon the background dose level.

As an example of the incremental potency argument we consider how the regulation of benzene might be modified if we do not assume a linear dose response. The limited data on leukemia in humans suggests (Lamm et al., 1989; Thorsland et al., 1988) that benzene carcinogenicity proceeds with a nonlinear dose-response mechanism with toxic effects manifested via metabolites. Also, benzene is not a direct mutagen, making linearity more unlikely. We here assume that the dose response is a cubic ($n = 3$), but also make the usual assumption that benzene produces leukemias in the same way that background leukemias are produced. The risk ratio, at the dose where the data on benzene are anchored, is given by $RR = 12$. Then the incremental potency (near the background dose) is approximately $3/5$ of the straight line potency (as shown in Figure 3). An argument similar to this has been given for the risk of leukemia from benzene by Thorsland et al. (1988). To the best of our knowledge, this was the first quantitative discussion using the Crump et al. (1976) idea.

In general, the ratio of incremental potency to straight line potency is on the order of one, when $(3 < RR < 30)$ and $(2 < n < 5)$, which can justify the present regulatory approach. However, to the best of our knowledge, this simple justification has never been given before.

However, the usual assumption may not be correct. There is some evidence that benzene leukemias are always preceded by pancytopenia, which does not occur naturally in the population to any appreciable extent. If this is the case, background leukemias must be produced by an entirely different mechanism than the benzene leukemias, and the Crump et al. (1976) argument does not apply. Then the ratio of the incremental potency at a typical environmental dose of 1/50 of where the straight line potency is anchored to that straight line potency becomes $3 \times (1/50)^3$, or about 1/800, (assuming a cubic dose-response). This emphasizes that it is important not only to measure the biological dose response, but to have some ideas about the biological mechanisms; in particular, whether the mechanisms are the same for effects of the pollutant and of the background.

In the following discussion we will further illustrate this argument by applying it to the following cases: (1) exposure to particulate air pollutants, cigarette smoke, and coal dust, causing a reduction in lung function which can lead to premature death; (2) exposure to DBCP, causing a reduction in sperm count which can increase incidence of male infertility; (3) exposure to radiation in utero, causing a reduction in IQ and hence increased incidence of mental deficiency among children.

In each of these cases, others have suggested that there be a linear dose-response relationship, but have not related their suggestion to the general principle. In each case, it seems plausible that there is a linear relationship between an environmental (typically low-dose) exposure and an intermediate biological response, which itself leads to an effect on health. If the biological model is "plausible", even though
possibly not completely understood, we argue that the calculated risk must be taken seriously (Evans, Ozkaynak, and Wilson, 1982; Hill, 1965). Indeed, if plausibility is taken to require that every step in the mechanism leading to the toxic effect be clearly understood, we will fail to describe (and hence, predict and prevent) many risks to health. However, we must ensure that a hypothesis is rooted in available data and knowledge, and be willing to modify our statement of plausibility if and when new data appear, or we could spend time and money in fruitless chases.

Statistical limitations make it usually impossible to determine directly the dose-response relationship at the low doses of interest. However, it may be possible to determine the linearity (or otherwise) of the intermediate biological response, and this analysis suggests that is a good place for biological attention.

APPLICATION TO NONCARCINOGENIC END POINTS

Reduced Lung Function from Air Pollution, Smoking, and Coal Dust

Air Pollution

We have known that air pollution is bad for health for hundreds of years (as discussed in Wilson et al., 1980). Particulate stack emissions contain a variety of trace metals, organic matter, and radionuclides. High levels of particulate air pollution have been associated with excess mortality for centuries. In 1661, Evelyn attributed the rising death rate and increased incidence of respiratory disease in London to the combustion of coal (Evelyn, 1969). Extremely high levels of air pollution were associated with considerable excess mortality in episodes in the Meuse Valley, Belgium in 1930 (Ficket, 1936), Donora, Pennsylvania in 1948 (Shenck, et al., 1949), and London in 1952 (British Ministry of Health, 1954). During the 1952 episode in London, an estimated 4,000 excess deaths occurred in the metropolitan area, and British Smoke measurement of particulates had a maximum 24-hour average of about 4,000 μg/m³ (Logan, 1953).

But the idea that effects on health might be linear with dose is relatively young. In 1971, Nishawaki et al., (1971) plotted the death rate in several Japanese cities and showed that it was linear with air pollution variables. In a more systematic study, Lave and Seskin (1970; 1971; 1977) found a correlation between death rates in certain air pollution regions in the U.S. and various measures of air pollution, and suggested that the relationship is linear. Problems with confounding variables were raised, and since Lave and Seskin were not professionals in public health, their ideas were widely discounted. Others (Evans et al., 1982; Lipton, 1977; Mendelsohn and Orcutt, 1979; Bozzo et al., 1979) verified that the effect of air pollution is indeed present in the data, and appears to be incremental down to the lowest doses studied (confirming the arithmetic of Lave and Seskin), and went on to demonstrate the statistical association in subsequent data sets. Although critics (Viren, 1978) have suggested alternative explanations, such as confounding effects of cigarette smoking or migration of elderly and dying people, these arguments are not quantitatively satisfactory (Wilson et al., 1980).
Schimmel and Greenberg (1972) suggested that air pollution acts as an additional stress on subjects with pre-existing respiratory or cardiovascular problems. Evans et al., (1982) discussed the plausibility of this general model, and developed a more specific model (first suggested by Speizer), which we develop further here. The crucial step (which admittedly is not justified, but which they and we suggest is plausible) is to assume that each exposure to air pollution (or cigarette smoke) permanently destroys a small part of the lung, and lung function (as measured by Forced Vital Capacity, or "FVC") declines. At first sight, this would seem to lead to a threshold type dose-response, since most people have adequate lung function. But detailed examination suggests otherwise. On average, lung function rises during infancy and adolescence, peaking at about age 25, and declines continuously thereafter. Most people have enough lung function to be viable until age 100 or so (when most will have died from other causes anyway); but for a few on the left tail of the distribution, lung function will be insufficient at an earlier age and premature death will result from one of a number of causes.

When proposed in 1982, this model had little or no effect upon regulatory authorities. But since then, extensive measurements of the effect of air pollution on health in several cities by Dockery, Schwartz, Pope and collaborators have been made using improved methods of accounting for confounding exposures. Pope, Schwartz, and Ransom (1992) observed that the relative risk of death increased monotonically with increasing levels of respiratory particulates (the PM_{10} fraction), and that the relationship held at levels well below the National Ambient Air Quality Standard of 150 μg/m^{3}. In addition, Schwartz and Dockery (1992), and more detailed studies of Dockery (Dockery et al., 1985; 1988) found an incremental decrease in lung function, across a wide range of concentrations, with no sign of a threshold. Furthermore, Pope et al. (1991) found that elevated PM_{10} levels of 150 μg/m^{3} were associated with an approximate 3 to 8% decline in lung function (as measured by Peak Expiratory Flow). It is now possible to make the model quantitative, extrapolating from the data on smoking, as examined below.

Other information adds plausibility to the overall argument. Many early workers on air pollution related effects on health to sulfur dioxides concentrations (Dockery, Schwartz, and Spengler, 1992). Over the years, this has been replaced by sulfate particulate levels, total suspended particulates (TSP), and more recently, respirable particulates (less than 10 micrometers in diameter - the PM_{10} fraction). These changes have been guided by a mix of common sense, animal data (Amdur, 1990), considerations of aerosol dynamics, and the recent direct data on health. Even so, it must be noted that our model does not directly address the immediate effects of air pollution found in the time series study. Further development will be necessary to do so.

In addition to the death rate correlations, even moderately elevated concentrations of particulate pollution have been associated with reductions in children's pulmonary function (Dockery et al., 1982) and increased risk of bronchitis and other respiratory illnesses (Ware et al., 1986). Therefore, there is now a large quantity of data about adverse effects of respirable particulates, together with a plausible model. We suggest that attention should be paid to examining this, and other, models and verifying whether or not the consequences are indeed correct. It
would be desirable to have data on reduction of lung function by air pollution, as there is for cigarette smoke. The present data are not statistically significant. Although the detailed biological processes relating decline in lung function to death and other problems is unknown, these mechanisms are not known for the effects of cigarette smoking either. Thus, we believe that there is now much more than the usual amount of information needed to take the linear dose–response relationship between air pollution and mortality seriously in regulatory and other matters.

**Smoking**

It is reasonable to assume that reduction in lung function is an intermediate step in the chain of events leading to mortality in a large fraction of cases. Smoking, known to reduce lung function, therefore becomes a convenient analog in attempting to understand the chronic effects of air pollution, since the chronic effects of smoking have been extensively studied (Mattson, Pollack, and Cullen, 1987). As shown in Figure 5, Ashley *et al.* (1975) confirmed the relationship between reduced FVC and increased mortality, for smokers examined in the Framingham Heart Study. Assuming that the reduction in lung function caused exposure to air pollution occurs via the same mechanism, or the same family of mechanisms as that through which the effects of smoking are manifested, we can expect adverse health effects in some portion of the population for any reduction in

![Graph showing the relationship between FVC and probability of death](image)

**Figure 5.** Smoothed probability of death in 10 years (measured as forced vital capacity) at Examination 5 of the Framingham Heart Study, 18 year follow-up (Ashley *et al.*, 1975).
lung function, and hence any level of air pollution. This assumption can be tested by
asking the extent to which the deaths from cigarette smoking and from air pollution
have similar proximate causes (heart attacks, cancer, etc.).

In the Framingham Heart Study, a cohort study of 5,209 men and women aged
30–62 at the beginning of the study in 1948, subjects were monitored biennially for
20 years. Mortality in the cohort varied over a four-fold range in inverse relation to
the biennially measured FVC (Kannel et al., 1980). Excess mortality associated with
low FVC was shown to be independent of age, sex, cigarette smoking, relative body
weight, and the major cardiovascular risk factors (Ashley et al., 1975; Kannel et al.,
1980). In general, FVC was an exceptionally good predictor of both long- and short-
term mortality, second only to age when compared to the major cardiovascular risk
factors. As shown in Figure 6, in both smokers and nonsmokers, FVC declines
steadily with age, and the mean values of FVC are consistently lower for smokers
than for nonsmokers. Longitudinally, cigarette smokers showed a more rapid decline
in FVC in 10 years than did nonsmokers. Overall, a striking relation of FVC to
mortality was noted, which is not accounted for by smoking habits. Figure 7 shows
a relationship between FVC (liters) and age in smokers and nonsmokers, as
predicted by the Ashley et al. (1975) model.

The nonsmoker plot is for a man who never smokes. Both groups show a decline
in FVC as a function of age, but smokers' lung function declines more sharply. The
association between FVC and mortality persists in nonsmokers, even after persons
with pulmonary disease or chest deformity on the initial examination were excluded.
The overall association of FVC with mortality persists even after age-adjustment, as
can be seen in Figure 8. It is plausible that the association between FVC and
mortality is due to its relationship to cigarette smoking for several reasons. First,
mean FVC was lower in cigarette smokers than in nonsmokers, and was lowest in
heavy smokers. Furthermore, FVC declined more rapidly with age in smokers than
in nonsmokers. Finally, the rate of decline in FVC was dose-related. Relative trend
statistics showed negative values at every age, indicating that FVC fell faster in those
who died than in those who survived. In conclusion, Kannel et al. (1980) suggest
that determination of FVC is an efficient way to select groups of persons otherwise
without symptoms who are nonetheless destined for premature death.

Dockery et al. (1982) noted that, clinically, only a small fraction of smokers
(approximately 10 to 15%) develop serious respiratory impairment (low FVC
values). This suggests that a group within the population of smokers is more
susceptible to the effects of cigarette smoking. Their model implies that those
subjects who entered adulthood with lung function on the low end of the
distribution of pulmonary function are at the highest risk of disability associated
with respiratory impairment from cigarette smoking.

Coal Dust

Dust has been known to cause respiratory problems for millennia, at least since
Romans recorded the effects on British tin miners. By 1979 it was estimated that
100,000 miners were suffering from legally identifiable “black lung disease” (coal
worker’s pneumoconiosis, or “CWP”). The U.S. Coal Mine Health and Safety Act
REGRESSION OF MORTALITY ON FVC
FRAMINGHAM HEART STUDY
10 YEAR MORTALITY FOLLOW-UP

<table>
<thead>
<tr>
<th>Women</th>
<th>Age at Exam 5</th>
<th>Deaths/At Risk</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FVC</td>
<td>(T-Value)</td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>.267</td>
<td>-1.44</td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>.322</td>
<td>-2.01</td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>.597</td>
<td>-4.07</td>
<td></td>
</tr>
<tr>
<td>Av. Over Age Groups</td>
<td>.419</td>
<td>-4.51</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Men</th>
<th>Age at Exam 5</th>
<th>Deaths/At Risk</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FVC</td>
<td>(T-Value)</td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>.399</td>
<td>-2.47</td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>.301</td>
<td>-1.98</td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>.449</td>
<td>-3.38</td>
<td></td>
</tr>
<tr>
<td>Av. Over Age Groups</td>
<td>.389</td>
<td>-4.57</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6. Regression of FVC vs. mortality; Framingham Heart Study (adapted from Attfield, 1992).

Figure 7. FVC (liters) vs. age for men in Framingham cohort. Notice that the rate of decrease in FVC is greater in smokers that in nonsmokers (adapted from Ashley et al., 1975)
Low-Dose Linearity

(U.S. Congress, 1969) required reduction of respirable dust exposures in coal mines to below 3.0 mg/m³ beginning January 1, 1970, and to below 2.0 mg/m³ by 1973. The decision to reduce the levels was driven by data regarding prevention of CWP, but the level was selected to include feasibility. Follow-up studies conducted to assess the efficacy of the standard have concluded that it does not adequately protect miners’ health. Attfield and Morrin (1992) estimate that between 2% and 12% of miners exposed to a 2 mg/m³ dust environment in bituminous coal mines would be expected to develop CWP (category 2 or higher) after a 40-year job history. Risks for anthracite miners appear to be even greater.

Several recent studies have correlated prevalence of CWP with coal dust exposure (Seixas et al., 1992; Attfield, 1992; and Attfield and Hodous, 1992). Attfield and Hodous (1992) used data on dust exposure to develop dose-response models for forced expiratory volume (FEV), forced vital capacity (FVC), and (FEV/FVC). Their primary modeling consisted of linear regressions of FEV, FVC, and (FEV/FVC) against age, height, smoking status, pack-years of cigarette smoking, and estimated cumulative dust exposure. Separate models were fitted by smoking group with interaction variables, to inquire into any possible synergistic effects between dust exposure and smoking. Results did not indicate that the combined effect of smoking and dust exposure is worse than the additive effect of each.

Attfield and Hodous (1992) concluded that, in general, it appears that dust and smoking acted similarly on ventilatory function in causing a general shift of the FEV distribution to lower values. Hurley and Souter (1986) suggested that a subgroup of

\[ \text{FVC} = a + b \times (\text{pack years}) \]

<table>
<thead>
<tr>
<th>Age</th>
<th>Number</th>
<th>Coefficient (b)</th>
<th>SD (b)</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>37-49</td>
<td>354</td>
<td>-0.030</td>
<td>0.024</td>
<td>-1.24</td>
</tr>
<tr>
<td>50-59</td>
<td>211</td>
<td>-0.040</td>
<td>0.022</td>
<td>-2.24</td>
</tr>
<tr>
<td>60-69</td>
<td>136</td>
<td>-0.040</td>
<td>0.022</td>
<td>-1.82</td>
</tr>
</tbody>
</table>

*adapted from Ashley et al., (1975) data for men

<table>
<thead>
<tr>
<th>Age</th>
<th>Nonsmoker</th>
<th>S.D. (NSm)</th>
<th>Smokers</th>
<th>S.D. (Sm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>37-49</td>
<td>4.47</td>
<td>0.81</td>
<td>4.4</td>
<td>0.75</td>
</tr>
<tr>
<td>50-59</td>
<td>3.99</td>
<td>0.72</td>
<td>3.94</td>
<td>0.76</td>
</tr>
<tr>
<td>60-70</td>
<td>3.65</td>
<td>0.73</td>
<td>3.44</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Figure 8. Framingham cohort regression of FVC vs. smoking (pack years) (adapted from Ashley et al., 1975)
miners exhibit particularly severe dust-related losses in FEV. This especially sensitive group may correspond to those members of the population whose innate lung function places them at the left extreme of the normal distribution of lung function.

We conclude, therefore, that whether or not there is a true linear relationship relating response to the combined effect of dust and other irritants, there is no level of incremental dust exposure below which there is no effect on health. Since our model is a population model and assumes that some people already will have inadequate function at old age, this conclusion might be modified if workers are pre-selected to have a high Forced Expiratory Volume (FEV).

Decrease in Intelligence Quotient Following Radiation Exposure

One of the most dramatic, yet controversial, observations from study of the survivors of Hiroshima and Nagasaki is a reduction in intelligence quotient (IQ) among children who were in utero at the time. The BEIR V (NAS, 1990) report states that, although data are insufficient to precisely define the shape of the dose-response curve, they are consistent with no threshold for the reduction in IQ when the brain is in its most sensitive stage of development (in fetuses between 8-16 weeks of gestational age). Many people argue that there is nonetheless a threshold for severe mental retardation. On the contrary, we here argue (following BEIR V and the arguments of the previous section) that a linear dose-response relationship of mental retardation is a likely consequence of a linearity of IQ with dose of ionizing radiation.

As shown in Figure 9, adapted from BEIR V (NAS, 1990) there is a continuous diminution of intelligence with increasing dose. Within the group exposed 8-15 weeks after conception, the regression of the intelligence test (Koga) score on absorbed dose is linear. The range of decrease in intelligence test score is between 21 and 29 points at one Gy.

The distribution of intelligence quotients in a population is a bell shaped smooth curve. Although we approximate it here as a normal distribution, all that really matters is that the distribution is smooth at the low IQ side. Severe mental retardation, defined as having an IQ less than 70, occurs naturally in about 1.3% of the population. The risk of severe mental retardation is about 43% with 1 Gy exposure. Those members of the population at the lower left end of the frequency distribution of IQs before irradiation are most likely to be irreparably damaged. Assuming a normal distribution of IQs, the expected impact of radiation exposure on the prevalence of mental retardation was computed by BEIR V, as the integral from negative infinity to 70 of the IQ function. Thus, BEIR V predicts a 4% chance of mental retardation (thus defined) per 0.1 Gy radiation, when the fetus is exposed at the critical age.

The meager statistics on mental retardation prevents a direct proof or disproof of a threshold for mental retardation. However, the somewhat better statistics on the intermediate mechanism (reduction of IQ) enables a plausible assignment linearity to be made indirectly. This is an example of a general result.
Increase in Male Infertility Following Exposure to DBCP

1,2-dibromo-3-chloropropane (DBCP) was used as a nematicide in California vineyards to kill the nematodes that destroyed the vines. It was erroneously believed that the DBCP would not reach groundwater in appreciable quantities, and therefore its use was considered benign to the environment. DBCP has now been detected in 2,113 out of 5,288 drinking water wells sampled in California. Its use as a nematicide is now banned. Although care was taken in its manufacture to avoid exposure to airborne dusts, it transpired that workers in a factory owned by Shell Oil were exposed dermally, and many men became infertile. Further measurement shows that men exposed to DBCP have a reduced sperm count. It was widely assumed that the infertility was an acute toxic phenomenon with a threshold-type dose-response relationship, and that therefore only the carcinogenicity of DBCP need be taken into account at low doses. Recently, however, Pease, Vandenburg, and Hooper, (1991) pointed out that this may not be true. Their argument is another special case of the incremental potency argument here discussed.

The biological mechanism appears to be a reduction in sperm count as a man is exposed to DBCP that leads to the infertility. This reduction has been noticed not only with the Shell workers, but also with workers in Florida (see Figure 10). In any population of men, there is a distribution of sperm count. We simplify the discussion of fertility here by further assuming that any man whose sperm count falls below a certain threshold level (as illustrated on Figure 10) becomes infertile. Since 15% of all men are infertile, some men in the control group already lie below this threshold,

![Figure 9. Fetuses exposed to ionizing radiation 8-15 weeks after conception show loss of IQ points that could be proportional to dose (adapted from NAS, 1990).](image-url)
but more men in the exposed group fall below this threshold. Clearly any reduction in sperm count that applies proportionately to the whole population will result in a corresponding reduction in fertility. By the general argument above, the reduction in fertility will be incremental with the reduction in sperm count and linear, at low doses, with dose of DBCP. Tests on rabbits show that they too suffer a reduction in sperm count on exposure to DBCP (Figure 11). The crucial question here is the form of the relationship relating reduction of sperm count to dose. The data are consistent both with either a linear or nonlinear relationship between reduction and dose, as well as with a threshold at about 1 ppm DBCP. We believe that it is plausible that it could be linear, and in what follows therefore, we assume it is. We emphasize that this is the crucial biological issue.

![Graph showing sperm count in DBCP manufacturing workers.](image)

**Figure 10.** Sperm count in DBCP manufacturing workers. Data from Meistrich and Brown discussed in Meistrich and Brown (1983).
Low-Dose Linearity

The dose-sperm count relationship should obviously be derived for man if at all possible, although data exist for rabbits, the interspecies comparisons for anything to do with the reproductive system may be different than those for cancer induction, about which there has been considerable scientific and regulatory discussion (Calabrese, 1983; Goodman and Wilson, 1991). The data, even on rabbits (Figure 11), are far from definitive, and we again suggest that determining the dose/sperm-count relationship may be far more productive than a direct attempt to examine increases in infertility at low doses.

Why should we not assume that any given chemical can reduce sperm count, and therefore increase infertility? Should we take a default position that any chemical will have an effect unless proven otherwise? For DBCP, there are data on the effect on sperm count. For most chemicals, there are no data. This raises the issue, should all chemicals be considered to have an infertility risk if there are no data? Or exhaustively tested to show that they do not? If not, how do we provide an incentive to acquire data? Once we realize that infertility is only one of many interesting noncarcinogenic end points, we predict an enormous amount of work ahead for regulators, and procedures must no doubt be modified to cope with it.

Figure 11. Sperm count in rabbits. Data from Rao et al., 1982.
Policy Implications

Historically, poisons have been considered to exhibit thresholds. Paracelsus' famous remark, "the dose makes the poison," has often been taken to suggest variability in the threshold or toxic dose of substances, rather than differences in their low-dose carcinogenic potency and hence differences in probability of an effect at a low dose. However, in the 20th century this began to change. The first suggestion of a linear dose-response relationship was probably due to Jeffrey Crowther (1924), who formulated a theory that radiation-induced cancers result when an ionized cell becomes malignant and proceeds to multiply. Probability of cell ionization is inherently linear, since it is a product of the frequency of cell bombardment. If this initial idea is modified to allow that only one in many million cells becomes malignant (i.e., that most damaged cells are actually repaired by the body's defenses), linearity is still maintained, because the fraction is constant with dose. (Linearity would not hold if malignancy were found to result from some critical amount of radiation discreetly overwhelming the body's defenses.) Since that time, physicists have tended to accept Crowther's suggestion that cancer incidence is linear with exposure to radiation, but biologists and physicians have often continued to believe in thresholds.

Dose-response linearity was first introduced into public policy by the International Commission on Radiation Protection in 1927, which proposed that no one be exposed to any radiation without expectation of some benefit. (It is likely that they were thinking primarily of genetic mutations.) The idea that directly-acting carcinogens might exhibit low-dose linearity, whereas poisons might not, was already mooted in the 1930s. The idea that even small doses of carcinogens can have effects, although correspondingly small, is inherent in the legislative mandate of the Delaney clause of the Food, Drug and Cosmetic Act (FDA, 1957), which disallows any trace of a carcinogen in a food product. Peto (1979) led the argument that a linear dose-response mechanism should be assumed for asbestos-induced cancers (for purposes of prudent public policy). Nonetheless, it is widely suggested that a carcinogen acting through an indirect mechanism might exhibit nonlinearity, and recently it has been suggested that asbestos is an example of such a case. For example, the Chief Inspector of Factories in the UK, musing about asbestos, asked (56 years ago), "Is it the asbestos or the asbestosis it causes that initiates the lung cancer?" (Merewether, 1938). Scientists have not conclusively agreed on the answer to this, although Hughes and Wein (1991) have produced evidence that asbestosis is a necessary precursor to lung cancer. Then the Crump argument would not apply, since asbestosis is not a natural condition.

The yearning for binary ways of coping with risks (yes, it is safe/no, it is dangerous) finds its way into scientific discourse. Thirty years ago, although carcinogens were suspected of displaying linear dose-response relationships, there were few chemicals known to be carcinogens. In this setting, the simplest solution seemed to be to ban those which were known (Coulston, 1977), and if one or two turned out to have net benefits for society, to cope with these exceptions separately. In recent years, with the availability of more sensitive detection instruments, it has been shown that about half of all chemicals tested in animals cause cancer at some
Low-Dose Linearity

dose. Also, some known carcinogens have been shown to have benefits for society (for example, saccharin or alar). So today the simple banning solution seems inappropriate. Our extension of this approach to noncarcinogenic end points using the argument of Guess et al. (1977) underscores the inadequacy of this approach to regulation. In our view, there is no alternative to openly discussing what level of risk is acceptable and possible.

Dose–response linearity for noncancer end points depends upon the existence of particularly sensitive persons in the population. In contrast, the simplest model for cancer is that most of the population is approximately equally sensitive, and a random stochastic effect determines which persons will be victims. This suggests at once a number of possible policy differences, based upon the idea that society has a duty to protect sensitive individuals. In particular, if there is low dose linearity for noncancer end points underscores the question of whether regulations are being set to protect each individual to the highest level possible, or to protect the entire population to some reasonable level becomes important. Several questions are raised:

1. Should allowable exposure levels for DBCP be based upon the risk to an average male (assuming a threshold for infertility), or for the especially sensitive male with a low sperm count? (Radiation regulations already discern between pregnant women and less sensitive groups.)

2. Should we have separate occupational dose levels for the sensitive and for the average worker? The present situation seems to vary. Trades unions have disliked the idea of allowing healthy workers to work with higher concentration levels of lead than more sensitive workers. Yet in the U.S. Coal Mine Health and Safety Act of 1969 (U.S. Congress, 1969), workers who are sensitized can be removed from areas of high concentration.

3. For radiation protection, exposure of pregnant women, fertile women in the time that they might become pregnant, and children (persons under 18) is discouraged or forbidden.

4. Is it legitimate for an employer to screen for sensitivity to toxins in advance? (Some employers screen for cigarette smoking or drug use, and this has been upheld by the courts). Separating society into the sensitive and unsensitive, or healthy and nonhealthy raises human rights problems, that we only touch upon here but they have been handled so far. Identification of the portion of the population most at risk to the effects of a particular toxin may incur new problems.

We noted above that the coefficient of low-dose linearity depends upon the background incidence of the end point being discussed. Background levels of some pollutants will be different in rural and urban areas, so that regulations cannot simply be based upon "unit risk factors," as is presently done by the USEPA for carcinogens.

Since it is economically and physically infeasible to identify and avoid exposure to all harmful agents, we must settle for protecting against the most harmful and the least beneficial. It is intrinsically distasteful to consider that the health or life of certain individuals will be sacrificed for economic gain, even though the sacrifice
would be random. It is much more comfortable to espouse the doctrine of zero risk, and call for biological models. But, as has been discussed widely (Coulston, 1977; Gorman, 1993), the first makes regulations unwieldy and impractical, and often diverts scarce regulatory funds to protect against relatively insignificant risks, while the second is ducking the issue, since zero risk is unattainable, and a complete biological model for every pollutant unrealistic. As a society, we must accept and live with the idea that a necessary part of health risk assessment involves weighing benefits against costs (Wilson and Crouch, 1982).

The number of situations where there is a linear dose-response relationship is likely to be much greater than given by carcinogenesis alone, and suggests a more careful analysis of what risks are acceptable. The USEPA, for example, has set regulations to protect individuals from low levels of carcinogen, and set an "acceptable" or "de minimis" level of risk of one in a million per lifetime. If the USEPA tries to regulate the many noncarcinogenic end points to the same standard, the system could well be choked. A reexamination may well be necessary.

The models which we develop above depend on there being a background incidence of the effect of the pollutant. They also incorporate the assumption that the same mechanism operates in producing the observed background effects and the effects of the pollutant. In cases where the background incidence is zero for a particular pollutant's effects, our argument would not apply.

The Risk of Radiation-Induced Cancers: Hormesis

It is amusing to return to the discussion that Crowther started in 1924, of radiation-induced cancers. There has been recent discussion about whether the number of cancers at low doses is above or below the assumed straight line, proportional, plot. A positive, favorable, effect (hormesis) has been suggested at low doses. A biological mechanism (repair) has been suggested to explain why the number is below the line. If there is such a favorable effect at low doses, there is then a nonmonotonic dose-response, and the argument becomes very complex. It would still be important to know whether radiogenic cancers are assumed to result from the same mechanism as naturally occurring cancers. If they do so result, there could be situations where hormesis could be biologically correct in the absence of background, but not true in real situations where the background had already brought the total dose to a region of positive slope of the dose-response curve. The argument of this paragraph could also apply to situations where hormesis has been proposed for chemical carcinogenesis.

Many authors, and in particular, Pauling, Golman, and Stewart, have tried to argue that radiation exposure is a necessary first (initiating) step in all cancers, in which case the argument of Crump et al. (1976) would apply. But this idea predicts too many radiation-related cancers, and in particular, too large a variation with natural background. The assumption of a common mechanism seems, therefore, no more proven than in the cases which we discussed above. Indeed, some measurements suggest that leukemias induced by radiation therapy have a different cell structure than other acute myelogenous leukemias. If true, acute myelogenous leukemia might display practical nonlinearity, as already suggested by recent data from RERF.
Low-Dose Linearity

In this radiation example, as in the others, the importance of comparing the biological mechanism for the process by which the pollutant acts with that by which background cancers occurs becomes evident. Yet this has rarely been the subject of discussion.

APPLICATIONS IN ABSURD AND OTHER SITUATIONS

The Risk of Overcoats

As in mathematical proofs (*reductio ad absurdum*), it is instructive to try to construct absurd situations where the concept is legitimately applied, but the conclusions seem absurd. From a study of these, a lot can be learned. Recently, Koshland (1993) argued in jest that there is a risk associated with wearing an overcoat while crossing a busy street. The heavy overcoat restricts activity, and thereby makes it more difficult for an elderly, not so nimble, person to avoid some hazard such as an oncoming automobile. It therefore increases the likelihood of death. It is evident that a simple model would show linearity both with the weight of the overcoat, and with the frequency of its use. Koshland (1993) used this absurd situation to decry the use of linear dose-response relationships. We believe this is wrong. Rather, the absurd situation should be used to explain the importance of comparing risk and benefits of simple remedial actions.

The risk is clearly small, and is considerably exceeded by the benefit (even in the same units of health effects) of wearing an overcoat in winter. Of course in the summer, or in a tropical climate, where the benefit of an overcoat is actually negative, no one wears an overcoat! This seems to us an example of how people instinctively make risk/benefit calculations, particularly where there is negligible risk, all the time.

The Risk of Aspirin Use

The FDA (1977) calculated the total impact on society of saccharin use, by taking the total amount consumed and multiplying by the carcinogenic potency as derived from animal data, to be 470 excess cancers per year. This calculation would be correct if the animal data at the highest dose indeed apply (which is doubtful), and the dose-response is indeed proportional (linear with no threshold). Chester Richmond (1980) showed how a similar calculation leads to absurd results in a discussion of deaths due to aspirin use. Each aspirin tablet initiates a small amount of bleeding. Multiplying this coefficient (amount of bleeding per aspirin tablet) by the total amount of aspirin sold, he derived a figure for the total gallons of blood lost due to aspirin use. Then he made the critical step of dividing this by the blood content of a single person. He finally reached the absurd conclusion that aspirin has caused tens of thousands of people to bleed to death! This could lead to a policy decision to ban aspirin. Of course, examination of this argument shows that the logically incorrect step was dividing by the blood content of a person, although this step would lead to a correct conclusion in the FDA calculation of the total number of cancers from saccharin.
Although Richmond's calculation demonstrates that the naïve use of the linear dose-response model as used by the FDA for saccharin carcinogenicity can lead to absurd results, more careful use of the model is appropriate. Thus any bleeding caused by aspirin can aggravate a bleeding ulcer, or cause one to bleed which otherwise would not do so. This could be formally demonstrated by taking the slope of the distribution of people with a history of ulcers. We would then find that aspirin use will exacerbate ulcers. Rather than banning aspirin, the public policy conclusion, already widely applied, is the recommendation that those with digestive upsets refrain from taking aspirin.

These “absurd” arguments suggest common sense approaches that society has adopted to these situations, and we suggest that these common sense approaches be extended to other situations. Indeed, in discussions with physicians we have found that they will accept the arguments of this paper for those people who already have ailments. But over age 65, most of us have some bodily restrictions, so this is a general problem.

CONCLUSIONS

We have argued that low-dose linearity may be a common situation for environmental toxins of interest to public health, rather than the rarity that has often been assumed. Any biological end point is likely to display this linearity, in proportion to the extent that this particular end point exists in a subsection of society. The public policy issues that are raised by this argument are inextricably linked to societal approaches to especially sensitive groups.

In fact, a precedent exists for the application of this type of “acceptable risk” approach. The Coal Mine Health and Safety Act (U.S. Congress, 1969) requires that each miner working at the coal face be tested once every five years for early signs of CWP. If any sign is detected, the miner is reassigned to a work site closer to the surface, where exposure to coal dust is lower. In this way, individuals with greater susceptibility to lung damage are identified and only allowed to be exposed to a level of acceptable risk. Clearly, this law recognizes the interindividual variability which leads to differences in response. This interindividual variability is what leads to there being a response at the population level for any level of exposure.

Terminology contributes to the traditional bias toward consideration of individual thresholds rather than a linearity for a group. When biologists discuss “low doses,” they usually mean low absolute doses, not low incremental doses. Whereas, when we speak of “low environmental doses,” we think of low incremental doses. Since our mission is to define a dose-response relationship relevant to environmental doses, attempts to understand the relationship at the origin of the dose-response curve (as pictured in Figure 1), may be irrelevant. Instead, we should strive to understand the relationship at the “background dose.”

The arguments in this paper should not be construed to imply that the authors reject biologically based models. On the contrary, we welcome biological models, and suggest, by means of the macroscopic arguments, where a good biological understanding can be most useful. A biological model is often taken to mean a full
Low-Dose Linearity

understanding of how a particular "end point," be it lesion or tumor, is produced by the agent (pollutant) at issue. In situations where we are discussing an "end point" that occurs naturally with considerable frequency, we suggest that an emphasis on the more limited questions of whether the pollutant acts in the same way as the background may be more tractable, and therefore more fruitful.

We also reopen the issue of the default assumptions for a regulatory procedure. Although it is usually hard to find a scientific reason for a particular regulatory assumption, the default assumption that a carcinogen has a linear dose-response relationship is often considered plausible. We argue that for many end points it is plausible that a pollutant acts in the same way as a background. Many scientists would consider it to be plausible (in the absence of contrary evidence) that a pollutant contributes to many end points. If plausibility is all that is required for regulation, then including noncancer end points in the regulatory scheme in this way could increase the regulatory load many-fold, and emphasizes the need for a logical way of prioritization.

*Acknowledgements*

The authors thank our many colleagues, and in particular, Dr. Mary Amdur, Dr. Douglas Dockery, and Dr. Jack Spengler, for numerous helpful discussions of dose-response relationships over the years. Dr. John Evans, in particular, has helped us extensively with the models of air pollution; suggested the example on IQ variation; and critically read this manuscript. Martha Crawford was supported by a fellowship from the National Science Foundation.

References


Crawford and Wilson


Low-Dose Linearity


