"Safety" Testing of Carcinogenic Agents[1]

NATHAN MANTEL *and* W. RAY BRYAN, *Biometry Branch and Laboratory of Viral Oncology, National Cancer Institute,*[2] *Bethesda, Maryland*

SUMMARY

The problem of determining what dose levels of an agent are safe, *e.g.*, non-carcinogenic, cannot be resolved unless one first defines some level of permissible risk, no matter how small, rather than insisting on absolute safety. Both because of practical considerations and statistical variation, the determination of low-risk dose levels, for example 1/100 million, cannot be made directly but must be by extrapolation from observed data. A conservative approach for doing so is given. In addition to an arbitrary definition of "virtual safety," it is necessary to define an arbitrarily high statistical assurance level and a rule for extrapolation by use of an arbitrarily shallow slope. Illustrative data by Bryan and Shimkin (J. Nat. Cancer Inst. 3: 503–531, 1943) on the carcinogenic action of methylcholanthrene yield a "safe," 1/100 million, dose of 9×10^{-4} mg per mouse when a statistical assurance level of 99 percent and a conservative probit slope of 1 normal deviate *per log for extrapolation* are used. The principles given are of general applicability in other safety-testing problems, the point of emphasis being that since direct observation cannot be made that the risk at some dose level is clearly low, indirect conservative procedures for the determination of low risk levels must be made. The arbitrary risks and definitions for so doing may change with circumstances. The procedure does not require specification of an experimental protocol; the "safe" dose is determined on the basis of whatever data are available. Minimum protocols may, however, be desirable, since greater amounts of data will ordinarily permit specifying large "safe" levels.—J. Nat. Cancer Inst. 27: 455–470, 1961.

ALTHOUGH IT is not definitely known whether all chemical compounds that induce cancer in experimental animals will also cause cancer in man, it has been fairly well established that, with the possible exception of arsenical compounds, every chemical or physical agent known to produce cancer in man will likewise do so in one or more species of lower animals (*1*).

Of necessity, the potential deleterious effects of chemical compounds must be tested in laboratory animals. The reaction of a particular species of animal does not constitute proof that humans will react similarly, but

[1] Received for publication March 20, 1961.

[2] National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare.

597873—61——20

...ention is drawn to the possibility ... urged to investigate its co... ight status.

why new animal models when vaccine data is available?

the only recourse the investigator has in implementing test programs for control of the human environment with respect to injurious chemical or physical agents is to proceed under the assumption that any substance harmful to animals is potentially harmful to man. Also in carrying out control tests in animals one must proceed, initially, as if protection of the animal population were the actual problem. When reliable estimates of the doses tolerated (with specified probability) are achieved for the animal population, the problem then becomes one of judgment, based on accumulative experience, in transferring the implications of the results to man. Results of testing with a variety of animals could suggest how extrapolation could be made properly to mammalian species of higher order or larger size.

Many chemical compounds may be harmful at certain concentrations, though beneficial at others. The control of "toxic" or "harmful" substances therefore does not imply the necessity of their complete or absolute elimination, which in some cases would be either impossible or economically infeasible, but their reduction to concentrations that can be tolerated by essentially all individuals of the population at risk.

In rapidly acting toxic substances that are either quickly eliminated from the body or are readily transformed to less harmful compounds through metabolic processes, the estimation of tolerated levels is not too difficult. With a reasonable allowance for an extra margin of safety introduced in the form of an arbitrary "safety factor," the results obtained in laboratory animals can be successfully projected to humans. For the most part, modern pharmacology is based on just such usage of laboratory animals.

There are other compounds, however, for which the results obtained in laboratory animals cannot be so confidently projected to man. These substances are not readily excreted or metabolized, and because of their weak solubility in aqueous solution, may remain in the body or on body surfaces for very long periods. Other substances may be metabolized to some extent, but because of their selective affinity for cells of certain types they may accumulate selectively within these cells to yield harmfully high concentrations if supplied on a continuing basis. Most chemical compounds that cause cancer, *e.g.*, polycyclic aromatic hydrocarbons, naphthylamines, azo dyes, and some steroids, have properties falling into one or the other of these categories.

Still other compounds have an immediate initial effect, which is not considered harmful, and are rapidly eliminated or metabolized; yet, during their brief sojourn in the body they produce some critical intracellular damage or change, which does not manifest itself until later. Urethan, which interferes with nucleic acid metabolism, is a classical example of a compound of this type. Injected parenterally, urethan acts immediately to induce transient anesthesia, but single anesthetic doses may cause lung tumors in mice many months later. Fortunately, this drug was never approved for use as an anesthetic for man.

Biometric methodology has been highly developed for the study of

acute toxicity and rapidly developing biological reactions of other types (*see 2–4* for review). Some progress has been made in the development of methods for the analysis of responses that are greatly prolonged in time, such as cancer (*see 5* for review); however, the problems associated with the estimation of tolerated levels of carcinogenic agents and other chronically acting compounds have not yet received adequate study.

The purpose of this communication is to discuss the major problems in the development of methods for estimating limits of tolerance, or "safety levels," of carcinogenic compounds and to describe certain biometric procedures, based on available data, that are applicable to presently known carcinogenic compounds and experimental animal systems. Factors which one must consider in transferring the inferences derived from results in animals to man are also discussed. The suggested procedures are not restricted to carcinogenic agents, but are applicable, in principle, to compounds which cause various other types of harmful reactions or disease.

SOME PROBLEMS IN PLANNING AND ANALYSIS OF SAFETY STUDIES

In a safety-testing program both the design of the experimental protocols and the method of analysis or interpretation of resulting data must be determined. The experimental design cannot be properly determined apart from the plan for analysis of the data.

Certain issues in a safety-testing program must be resolved in advance. These relate to what we mean by safety and what kind of feasible results we are willing to accept as proof of safety. Settling of these issues, if necessary on some arbitrary but conservative basis, may permit answers to problems that would otherwise be insoluble. These problems are:

1) *How safe is safe?* Absolute safety can never be unquestionably demonstrated experimentally. Rather, experimental results can be used only to establish limits on the risk involved. With the specification of some level of risk, no matter how small, the possibility of determining whether or not that risk is exceeded opens. We may, for example, assume that a risk of 1/100 million is so low as to constitute "virtual safety." Other arbitrary definitions of "virtual safety" may be employed as conditions require.

Incidentally, an inflexible requirement for *absolute safety* may lead to acceptance of high levels of hazard. The impossibility of really demonstrating absolute safety leads to the acceptance, as a satisfactory demonstration, that no hazard was observed in an experimental protocol of moderately large size, 100 or even 1,000 animals. Such evidence, however, only provides assurance, at the 99 percent probability level, that the true risk is under 4.5 percent in the 100 animals or 0.46 percent in the 1,000 animals.

2) *What constitutes proof of safety?* In principle, one could use an experimental protocol sufficiently large to demonstrate that "virtual

safety" obtained. For this purpose it must be realized that an observed outcome of no tumors among 100 million treated mice does not necessarily demonstrate clearly that treatment was either absolutely or even virtually safe. This outcome could arise with a probability of 1 percent, even if the risk involved were as high as 4.6/100 million. It would in fact require a total of some 460 million tumor-free mice to demonstrate at the 99 percent assurance level that "virtual safety" obtained. Similarly, tumor-free results for 10,000 mice would only indicate that the risk was less than 1/2,200 and it would require tumor-free results in a total of some 450 mice to establish with high probability that the risk was under 1 percent. Studies of feasible size can be used to establish directly only risks of the order of 1/100 or higher. Data from such studies can be used to ascertain the treatment level consonant with a prescribed risk or to establish limits on the risk for a particular treatment level. The determination of "safe" levels can be made only by indirect methods extrapolating from the data obtained in a feasible study.

The use of extremely large studies to establish safety may well be self-defeating. The almost certain occurrence of unusual syndromes in one or more of a large number of test animals, albeit these may have arisen spontaneously, will require admitting the possibility that they may be attributable to drug treatment.

3) *How can protocol data be extrapolated safely?* Since it is only feasible to use experimental protocols for the direct determination of relatively high-risk dose levels, *e.g.*, 1 percent, which makes extrapolation methods necessary, one must consider that such extrapolation methods might yield misleading results. Procedures exist which permit extrapolating the results obtained at a number of test-agent levels to determine the dose level corresponding to any desired degree of risk and to establish, with a high level of assurance, a minimum bound on this dose level. These methods, however, are based on the assumption that the relationship observed between tumor occurrence and dose at the levels tested will continue to apply in the regions to which extrapolation is being made. The validity of such an assumption cannot be tested and, if it is false, may lead to a serious overestimate of the "safe" level. Such overestimation would arise if the relationship of response to dose is less pronounced at the dose levels to which extrapolation is made than at the levels at which tests were performed. Another related source of difficulty is that tests are performed on relatively pure inbred strains of laboratory animals. Characteristically, such pure strains will show steep dose-response relationships, while the heterogeneous population to which it is intended to apply the results of testing may exhibit a shallow relationship.

To avoid the risk of overestimation of "safe" levels which may result from extrapolating with too steep a slope, it is suggested here that a conservative result may be obtained by extrapolation with an arbitrarily low slope from the data at hand. For example, quantal-response data, that is, all-or-none response, frequently exhibit a somewhat linear relationship when plotted on probability paper with a normal or probit

scaling for the percent responding and a logarithmic scaling on dose. According to the kind of system being investigated the dose-response slopes observed may vary widely. With systemic poisons, rather steep dose-response slopes are generally obtained, reflecting the narrow logarithmic range between the lowest dose levels at which any toxic deaths occur and the levels which are lethal for all animals—such response slopes are on the order of 10 to 50 or more probits per common logarithmic or tenfold dose increase (the technical definition of this slope is not given here; it permits one to perform any necessary extrapolations). The all-or-none response also arises in the study of the therapeutic effects of antibiotics. The response slopes in these instances are generally shallower, on the order of 3 probits per common logarithm. Lower slopes on the order of 2 do arise in virus-assay work, but in these instances it may be that use of a different response curve, the single-hit or one-particle curve, would be more appropriate. From such experience it would appear that the use for purposes of extrapolation of a slope as low as one probit per common logarithm is likely to be conservative. While slopes in the regions to which we wish to extrapolate cannot be established, the suggested slope of one is rather low compared with that ordinarily obtained in the observable region.

The indicated low slope is a key feature in the method to be suggested as conservative for the establishment of "safe" levels. For this reason it may be well to make clear just how weak or strong are the assumptions being made in its use. In fact the only assumption being made for the procedure to be conservative is that whatever form the true response curve may take over the region of extrapolation, the average slope is not less than the assumed one. There is no requirement that the true response curve be linear or even that the true slope should nowhere be less than assumed. The use of the indicated conservative slope is, of course, arbitrary. Other values may be specified and other scales for extrapolation may be employed.

Once answers to the three questions are provided, a defined level of "virtual safety," a prescribed level of statistical assurance, and a conservative rule for extrapolation, it becomes possible to determine, from protocol data, "safe" dose levels and to undertake the planning of any necessary experimental protocols. This is considered in the succeeding sections.

ANALYSIS OF RESULTS AT A SINGLE DOSE LEVEL

In what follows we will take the defined level of "virtual safety" to be 1/100 million, the statistical assurance level to be 99 percent. Extrapolation will be on the basis of 1 normal deviate or probit per common log or tenfold change in dose.

To illustrate how definitive "safe" levels are obtained, consider that a prescribed dose of an agent has elicited no tumors in a group of 100 ex-

perimental animals. While the observed rate of tumor occurrence is 0 percent, we will take, as an upper limit on the true rate, that risk for which the probability for occurrence of as few as zero tumors is 1 percent (100% less the assurance level of 99%). This is given by the solution for P to the equation

$$(1 - P)^{100} = 0.01.$$

Solving, we have

$$100 \log (1 - P) = \log 0.01 = -2; \log (1 - P) = -0.02 = 9.98\text{-}10;$$
$$1 - P = 0.955; P = 0.045 \text{ or } 4.5 \text{ percent.}$$

We now know that the observed outcome of no tumors in 100 animals is consistent with the possibility that the true risk was, in fact, 4.5 percent. From tables of the normal probability function (*6*) we determine that the normal deviate, Y, such that the integral from $-\infty$ to Y equals 0.045 is -1.695, for a probit value of $3.305 = 5 - 1.695$. However, the normal deviate, Y_0, corresponding to a risk of 1/100 million can similarly be determined as -5.612, the probit being -0.612. The upper limit on the risk for the dose employed is $3.917 = -1.695 - (-5.612)$ normal deviates above the desired safe risk, and, at a slope of one normal deviate per common log, it is necessary to reduce the log dose by 3.917 logs to attain a "safe" level. The antilog of 3.917 being about 8,300, it is determined that the "safe" dose is 1/8,300 times that which had been tested.

Table 1 shows the preceding results together with those for several other hypothetical experiment sizes in which no tumors were observed to occur in a group of treated mice. For each such group the table shows the greatest risk consistent, at the 99 percent assurance level, with the observed outcome. Also shown are the corresponding conservative estimates, with the slope of one normal deviate per tenfold dose increase, of the "safe" (1/100 million) dose level expressed as a fraction of the dosage tested. The larger the experimental group among which no tumors occurred, the greater is the value determined as the "safe" dose.

TABLE 1.—Illustration of "safe" doses determined when no risk was observed in single groups

Number of mice with tumors/ No. tested	0/10	0/50	0/100	0/500	0/1,000
Upper limit on tumor risk at level employed, 99 percent assurance (percent)	37	8.8	4.5	0.92	0.45
Estimated "safe" dose (1/100 million) dose employed = 1	1/190,000	1/18,000	1/8,300	1/1,800	1/1,000

Study of this table indicates that a control system can be established without the need for specifying a design protocol, though there might be some merit in specifying a minimum size. For, whether the amount of evidence adduced to show that an agent is safe is great or small, it can be properly weighted to determine conservative safe limits. If the promulgator of a drug wishes to have high tolerances established for his

compound, it would be worth while for him to produce results of experiments with a large number of animals, which would show that the agent is not especially dangerous at certain dose levels. Where the lack of danger is demonstrated with a small number of animals or as a result of testing at rather low doses, the dose levels determined as "safe" will be lower. A control system can be constructed about the possibility for interpreting the data submitted, with no specification needed as to how much data should be obtained.

THE CASE OF SOME OBSERVED RISK

The method indicated for determining "safe" levels is not restricted to the case in which no observable danger was noted. That case was used for illustration because of the simpler mathematical solution involved. In general it will be that an experiment testing n animals will yield r unfavorable results (tumors). The upper limit on risk at the 99 percent assurance level is then the solution for P to

$$\sum_{i=0}^{r} nC_i P^i (1-P)^{n-i} = 0.01$$

or

$$\sum_{i=r+1}^{n} nC_i P^i (1-P)^{n-i} = 0.99$$

The solution for P is a value such that the chance of observing as few or fewer than r tumors is 1 percent. At values for P in excess of the solution, the chance for such an outcome is less than 1 percent.

The preceding equation can be solved approximately by reference to tables of cumulative binomial probabilities. Such tables, as for examples those of the Ordnance Corps (7), show values of quantities such as

$$\sum_{i=r+1}^{n} nC_i P^i (1-P)^{n-i}.$$

Let us consider as a hypothetical outcome that, of $n = 100$ mice, $r = 10$ have developed tumors, for an observed rate of 10 percent. Referring to these tables we see that for $P = 0.19$

$$\sum_{i=11}^{100} nC_i P^i (1-P)^{n-i} = 0.9891$$

and that for $P = 0.20$

$$\sum_{i=11}^{100} nC_i P^i (1-P)^{n-i} = 0.9943.$$

We may take the solution for P as approximately 0.192. The calculating procedure follows as before. The normal deviate corresponding to a 19:2 percent probability is −0.871 and, at the slope assumed, it will require a reduction in log dose of 4.741 = −0.871−(−5.612) to obtain a risk of 1/100 million. The "safe" dose is then determined as 1/55,000 of the dose which had been employed, 55,000 being the antilog of 4.741.

USE OF CONTROL DATA

In the methodology shown, it was assumed that the response of interest, appearance of tumors, did not occur spontaneously. In general, however, it will be desirable to use controls to check this and to allow data obtained for such controls to modify the determination of "safe" dose made. However, with the method already shown, failure to use controls or to take control data into account will result in more conservative determinations of the "safe" dose. If, in fact, spontaneous rates are rather low, they will have little effect on the determination made. For this reason we may adopt a procedure which is somewhat more conservative than necessary for taking control data into account. This we can do by taking, as before, the upper limit for the risk in the treated group as the solution for P_t to

$$\sum_{i=0}^{r_t} n_t C_i P_t^i (1-P_t)^{n_t-i} = 0.01$$

while the lower limit on the control group risk is the solution for P_c to

$$\sum_{i=r_c}^{n_c} n_c C_i P_c^i (1-P_c)^{n_c-i} = 0.01$$

where r_t of n_t treated animals and r_c of n_c control animals, respectively, showed positive response. These equations can be solved through the use of binomial tables.

At this point Abbott's formula (*8*) can be used to obtain a modified value for the treated-group risk, this being computed as

$$P_t' = (P_t - P_c)/(1-P_c).$$

The computation follows as before, the normal deviate being obtained corresponding to P_t'.[1] Since P_t' cannot exceed P_t, the use of control data cannot result in decreased values for the calculated "safe" dose.

ANALYSIS OF RESULTS OBTAINED AT SEVERAL DOSE LEVELS

When an agent is tested, it is sometimes desirable to do so over a number, perhaps even a wide range, of dose levels. In such instance, all the available data should be considered in the determination of "safe" levels.

[1] The statistician may object to the moderately conservative procedure described here for determining P_t'. A more rigorous procedure would require setting limits on the ratio of binomial parameters. P_t' would be given by reducing by unity the upper limit on the ratio of control-to-treatment nonresponse probabilities.

As already indicat
observed response sh
variation of the fitt
methods conservativ
large study sizes. I
the result that extr
the slope obtaining
sizes or studies, the
slope, taking accour
may result in extr
lower confidence li
than that which w
is used for extrapol

An alternative d
example, fitting th
determine the lowe
moderate percenta
with the observed
conservative value
1 percent dosage, e

While the proce
is based on the vali
results have derive
But for bioassay p
exactly appropriat
ably good job of g
can yield reasonab
ations being comp
use of an invalid
of dose levels cor

In many situati
in others quite s
large experiment
reflected in narro
to risks in other
one's estimate of
is violated.

Accordingly, w
we will consider
single dose-level
The only assump
tive and it is, of
instances in whic
is not misleading
generally appro
section.

Parametric Procedures

As already indicated, it may be unwise to extrapolate the data with the observed response slope. Procedures for taking into account the statistical variation of the fitted slope will not suffice to make such extrapolation methods conservative. To see this, one need only consider the use of quite large study sizes. In this case, statistical variation will be negligible with the result that extrapolation to low-risk levels will be substantially with the slope obtaining in the observable range. (However, with small study sizes or *studies*, the designs of which are inefficient for estimation of the slope, taking account of the statistical variation of the slope determination may result in extremely conservative *estimates* of "safe" levels. The lower confidence limit on the "safe" dose in such instances may be less than that which would be obtained when an arbitrarily low slope value is used for extrapolation, as suggested.)

An alternative device could be to employ parametric procedures, for example, fitting the maximum likelihood probit line to the data (*9*), to determine the lower confidence limit on *the dose corresponding* to some moderate percentage risk, *e.g.*, 1 percent, to which risk-level extrapolation with the observed slope is considered reliable. Then, using an arbitrarily conservative value for the slope, and anchoring at the lower limit on the 1 percent dosage, one could extrapolate to the desired "safe" level.

While the procedure just indicated is straightforward, its employment is based on the validity of the parametric function employed. Quite useful results have derived from such parametric assumptions in bioassay work. But for bioassay purposes it is not essential that a parametric function be exactly appropriate; it is sufficient that the function assumed do a reasonably good job of graduation. (Even somewhat inappropriate curve forms can yield reasonably good relative potency estimates as long as the preparations being compared are tested over the same regions of response.) The use of an invalid parametric function can lead to inappropriate estimates of dose levels corresponding even to moderate risks.

In many situations the estimates may be only moderately inappropriate, in others quite serious. With parametric procedures, the use of rather large experimental groups in one region of the response curve can be reflected in narrow-range confidence intervals for dose levels corresponding to risks in other ranges. This can produce a false sense of security in one's estimate of the moderate-risk dose level when the parametric model is violated.

Accordingly, while agreeing that parametric procedures may be useful, we will consider the possibility for extending the method described for the single dose-level case without the need for assuming any particular model. The only assumption is that the arbitrary low slope assumed is conservative and it is, of course, implicit that the response curve is monotone. In instances in which it can be recognized that use of a parametric procedure is not misleading, workers may prefer this procedure rather than the more generally appropriate nonparametric procedures described in the next section.

Nonparametric Procedures

In the preceding section it was suggested that the use of parametric methods, while straightforward, could lead to nonconservative results. There are no simple fixed rules for conservative estimates of the "safe" dose when several dose levels are employed and one is unwilling to make assumptions about the dose-response curve in the region of observation. How estimates can be made in these circumstances can best be demonstrated by illustration.[4]

We will begin with some simple ideas. Suppose investigators at two laboratories independently test an agent at a level of 100 mg/kg. At the first, with 500 mice tested, no tumors are observed, and with the methods described previously the "safe" dose is estimated as 100 mg/kg/1,800 = 0.056 mg/kg (*cf.* table 1). A somewhat lower "safe" dose of 0.012 mg/kg is obtained at the second laboratory, based on the observation of no tumors among only 100 mice. It can readily be recognized here that it would be inappropriate to reject the high "safe" level of the first laboratory just because of the low estimate obtained at the second laboratory. The two sets of data are consistent and in fact confirm each other. If any modification is to be made, it should be to consider that, with results combined, no tumors have occurred among 600 mice which would lead to a safe dose of about 0.065 mg/kg.

Suppose that at still a third laboratory, tests are made at a dose of 50 mg/kg and, with no tumors occurring among 500 mice, the calculated "safe" dose at that laboratory is 0.028 mg/kg. Here again we can see that the "safe" dose obtained at the first laboratory should not be modified downward just because a consistent result at the third laboratory yielded a lower "safe" dose. If anything, it should be considered that the 500 mice not responding at the higher dose at laboratory 1 would not have responded at the lower dose employed at laboratory 3. With these 500 mice treated as nonresponders at the low dose, there is then, including those at laboratory 3, a total of 1,000 mice not responding at the low dose. (Laboratory 2 results are being ignored for this illustration.) This yields as a calculated "safe" dose 50 mg/kg/1,000 = 0.050 mg/kg. In the present case the "safe" dose based on the combined calculation is less than that for the data of laboratory 1 alone of 0.056 mg/kg and so the higher figure is retained. Had the combined calculation led to a higher "safe" dose it would have been correct to take that as the estimate.

The point of these illustrations is that, when the data obtained from a series of doses are consistent with each other, it is appropriate to take as the calculated "safe" level the highest one pertaining to the results at any one dose. Even a higher "safe" level may be taken when it can be obtained through a justifiable combination of the results at the various doses used. What is meant here by "justifiable" combinations can be seen from the following example in which hypothetical results at four dose levels, low, middle, and high, are considered.

[4] An alternative method to the one about to be described is given as an appendix.

urged to investigate

Dose in size order	Observed results Number of tumors/ number of mice	"Justifiable" combined results Combined number of tumors/combined number of mice			
1	0/100	(0/100);[a]	0/200;	1/300;	5/400
2	0/100	0/100;	1/200;	5/300	
3	1/100	1/100;	5/200		
4	4/100	4/100			

[a] Parentheses indicate that this result need not be considered, as the next must yield a higher value for the "safe" dose.

In the absence of inversions in the data, we can determine the "justifiable" combined results at a dose by adding to the results at that dose, in succession, the results at still higher doses. A calculated "safe" dose can be determined for each dose used and for each of the various "justifiable" combined results corresponding to each dose. The over-all "safe" dose would be the highest of the various determinations.

Where data show an inversion the procedure is altered. Consider a simple example:

Dose in size order	Observed results Number of tumors/ number of mice	"Justifiable" combined results Combined No. of tumors/combined No. of mice
1	1/100	(1/100);1/200
2	0/100	[(0/100);1/100]

At the first dose level it is clear that the calculated "safe" dose would be larger if based on the combined results for both levels than if based on the results observed at this level alone; accordingly, as noted in one instance in the preceding example, the result at the lower level alone is shown in parentheses. At the higher dose level an inversion occurs; there is a lower incidence of tumors even though the dose level is higher. In view of the inversion, one would be less willing to accept as "safe" the calculated value obtained on the basis of results for this dose level alone. The two alternative results shown in brackets at this dose level are the results at this dose level and the contradictory results at the lower dose level. The significance of the use of brackets here is that the calculated "safe" value is now to be taken as the lesser of the values suggested by the alternative results. In the present instance, the result at the lower dose 1/100 would yield the lower "safe" dose and so the alternative result is shown in parentheses. (It will not always be necessarily true, when an inversion occurs, that the retained result will correspond to the higher tumor incidence at the lower dose level.) In the present example the calculated "safe" dose will be that corresponding to the first dose with combined result 1/200 or that corresponding to the second dose with retained result 1/100, whichever is the greater.

In practice the application of the methods just indicated is much simpler than the explanation would suggest. Ordinarily only one or perhaps two of the combined results at a dose level will need to be considered.

The results at some dose levels may immediately permit us to drop them from consideration. After only a limited amount of experience it should be possible so do this rather rapidly. The calculations are actually simpler than those for the maximum likelihood probit method. [In fact, the confidence limit procedures ordinarily employed in connection with the probit method are not fully satisfactory. A more appropriate method is described by Mantel and Patwary (*10*), but it could require a somewhat extravagant level of computational effort.] And, while for completeness, we have indicated the need for considering the possibility of inversions, this will ordinarily not pose a problem.

An Illustrative Example

An example from the literature shows how the procedure just discussed can be applied. The data, from Bryan and Shimkin (*11*), are the results obtained after a single injection of methylcholanthrene into mice, 12 different dose levels being used in the study. The reader may refer to the original article for details.

No peculiarities arise in this example. There are no inversions. At the four lowest levels no tumors occurred and the appropriate combined result is readily recognized in these instances. At the middle four levels it can be recognized that there is no point in combining results and, finally, the four highest levels can be disregarded as these all yielded 100 percent tumor occurrence.

The procedure is illustrated in table 2. The first three columns show, respectively, the dose, log dose, and the observed result. Column 4 shows each combined result considered, and there should be a separate line for each such result. In the present instance only one combined result required to be considered at each dose. For each such result, column 5 shows the calculated maximum risk at the 99 percent assurance level. These were obtained from binomial tables or calculated directly for the case of no tumors occurring. The normal deviate corresponding to the maximum risk is obtained from tables of the normal distribution and is shown in column 6. Finally, column 7 shows the calculated "safe" (1/100 million) log dose. The maximum for this, 2.962–10, appears in the second line, and the over-all calculated "safe" dose is 9×10^{-8} mg per mouse.

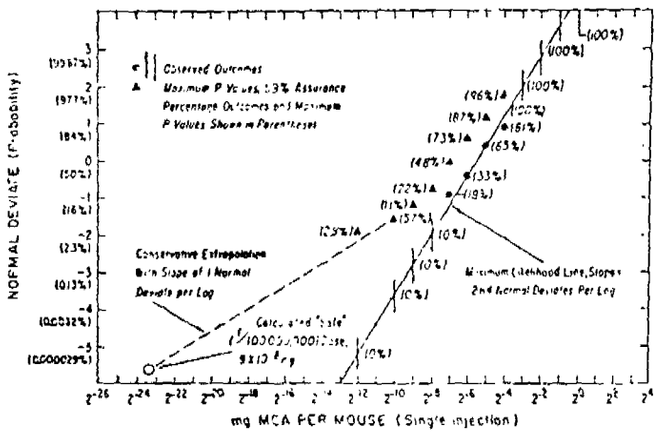
One might remark that this "safe" dose is so low as to make impractical any use of it which may result in its ingestion by humans. But we are dealing here with a rather potent carcinogen and if any compounds are to be assigned tolerated levels which are virtually zero, this is one of them.

Text-figure 1 shows graphically the results and analysis of the experiment just considered. Normal deviates are shown on the vertical scale, while on the horizontal scale the dose employed is shown as negative descending powers of 2. The points shown represent the outcomes at each dose level; 0 and 100 percent outcomes are shown by arrow. The solid line on the figure is the maximum likelihood probit line fitted to the

TABLE 2.—Illustration of methodology for determining the "safe" dose from results at several dose levels; data from Bryan and Shimkin (*11*)

Dose (mg/mouse)	Log dose	Result: No. of tumors / No. of mice	Combined result: No. of tumors / No. of mice	Maximum P value 99% assurance	Corresponding normal deviate	Calculated "safe" (1/100 million) log dose (2) − (6) − 5.612
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0.000214	6.388–10	0/79	0/158	0.0288	−1.899	2.675–10
0.000475	6.990–10	0/41	0/79	0.0566	−1.584	*2.962–10*
0.00195	7.291–10	0/19	0/38	0.1141	−1.205	2.884–10
0.0039	7.592–10	0/19	0/19	0.2152	−0.789	2.769–10
0.0078	7.893–10	3/17	3/17	0.480	−0.050	2.331–10
0.0156	8.194–10	6/18	6/18	0.729	+0.610	1.972–10
0.0312	8.495–10	13/20	13/20	0.871	+1.131	1.752–10
0.0625	8.796–10	17/21	17/21	0.958	+1.723	1.456–10
0.125	9.097–10	21/21	—	—	—	—
0.25	9.398–10	21/21	—	—	—	—
0.50	9.699–10	21/21	—	—	—	—
1.0	10.000–10	20/20	—	—	—	—

data. Above the first 8 data points the triangles shown correspond to the maximum P values of table 1. Extrapolation, with the slope of one normal deviate per common log to the over-all calculated "safe" value, is indicated by a broken line. All triangles, other than the one from which



TEXT-FIGURE 1.—Estimation of the "safe" dose from test results with a carcinogen methylcholanthrene, at several dose levels. At each test level both the observed percentage response and an upper limit, 99 percent assurance, based on combined data are shown. *Solid line* is the maximum likelihood probit line fitted to the data. The "safe" level of 9×10^{-8} mg per mouse is in this instance estimated by extrapolation with the conservative slope of 1 normal deviate per log from the upper limit on P at the second dose level.

extrapolation was made, should fall above the line. The difference in slope between the solid and the broken line may be noted.

RELATION TO OTHER SYSTEMS

While at various points in the preceding section the possibility for arbitrary selection of the assurance level, the level of safety desired, the slope value used for extrapolation, and even the extrapolation curve (one could use some scale other than normal deviates or probits along which a risk probability may be defined) have been emphasized, this should not be taken to mean that the methods suggested are completely general. Rather, in each case the nature of the risk situation should be considered.

Some principles do carry over, however. For example, one may be interested in trying to extinguish a bacterial or viral population by exposure to increasing temperatures or by increasingly long exposure to bactericidal or viricidal conditions. Determining the appropriate temperature or duration of exposure through the use of a conservatively shallow slope may be appropriate in such cases.

Consideration of the "single-particle" or "one-hit" problem (*12*) gives rise to a quite different answer than that developed in the preceding section. In this problem it is considered that a single particle can cause infection or death. If particles are distributed at random in material being inoculated, so that if there is an average of m particles per inoculation, the probability that a particular inoculation will contain none and so be safe is e^{-m}; the probability that it will not be safe is then $1-e^{-m}$. The problem is to determine from test data what reduction in the inoculum is necessary to ensure safety.

One can use directly the methodology already given to set maximum values on the risk as a result of the outcome of testing. This in turn establishes a maximum value for m as $-\log_e$(1-max. P value), the necessary reduction to any desired risk level following directly.

In this instance the inoculum corresponding to a risk of 1/100 million is approximately one millionth that corresponding to a risk of 1 percent. This contrasts sharply with the ratio of about 1/2,000 when extrapolation is made between these two risk levels with a slope of one probit per common log. While it has been shown that in the central region the "one-particle" curve mimics the probit curve with a slope of 2 probits per common log (*13*), the contrast would suggest that the comparatively steep slope of 2 noted virtually disappears in the lower tail. What further suggests itself is that the "single-particle" curve provides a most conservative rule for extrapolation. Where there is any suspicion that this curve may apply, the procedure described in the text should not be used.

One might also visualize a two-hit or two-stage process, the low-level probabilities for each stage being approximately proportional to the dose used. A reduction in dose by a factor of 1,000 should reduce the joint probability by a factor of 1 million. In this case the 1/100 million risk dosage would be 1/1,000th that of the 1 percent risk dose.

APPENDIX

Alternative Method

An alternative method to that described for handling data at several dose levels may be more appealing to the biometrically oriented reader. It was evolved in discussions with a reviewer (J. Cornfield) and is objective and intuitively satisfying.

In this method the data are considered to consist of a series of assays: one comprising results at the lowest dose level only; another comprising results at the two lowest levels; still another in which the results are given for those at the three lowest levels, etc. (Levels where there is interference, as for example lethality interfering with a carcinogenic response, should not be included.) For each assay, with appropriate probit procedures determine the maximum likelihood, 1/100 million dose level and its lower limit, subject to the restriction that the probit slope is known to be unity. Of all the lower limits on the 1/100 million dose level determined for the series of assays, the largest is selected as the "safe" dose.

A drawback to this procedure could be that standard methodology does not correctly permit the determination of lower limits on the 1/100 million level. Such methodology goes awry if, for example, there are no positive responses at the three lowest dose levels. A principle described by Mantel and Petwary (*10*) permits a solution. Calculate the chi-square goodness-of-fit, χ_1^2, (or log likelihood for those so inclined) for the maximum likelihood solution. (This chi square would be taken as zero if there are no positive responses at any of the dose levels considered; the maximum likelihood 1/100 million level is actually infinite. A zero chi square obtains also when only results at the lowest dose level are considered.) Consider alternative trial values for the 1/100 million dose level less than the maximum likelihood estimate. Each such trial value, in conjunction with the assumed slope of unity, provides an estimate of the response rate at each dose level of the assay. There is thus, in turn, a chi-square value, χ_2^2, for departures of the observed responses from the estimated responses based on the trial 1/100 million dose level. χ_2^2 will exceed χ_1^2 and for some trial value $\chi_2^2 - \chi_1^2$ will equal 5.412, the 98 percent upper limit on a chi square with one degree of freedom (a 98% limit on chi square corresponds to a 99% on the "safe" level).

At this point we introduce a modification which, for simplicity, has just been glossed over. Instead of computing χ_1^2 based on weights implied by the maximum likelihood solution, one should compute the quantity χ_1^{2*} based on departures from the maximum likelihood fit, the weightings being derived from the fit to the trial "safe" dose.

To illustrate: Suppose that at the i'th dose level, r_i of n_i animals respond; that the maximum likelihood estimate of the response rate is $\hat{P}_i$ and the estimate corresponding to the trial value considered is $\hat{P}_i^*$. Then

$$\chi_2^2 = \sum_i (r_i - n_i\hat{P}_i^*)^2/n_i\hat{P}_i^*(1 - \hat{P}_i^*)$$

and

$$\chi_1^{2*} = \sum_i (r_i - n_i\hat{P}_i)^2/n_i\hat{P}_i^*(1 - \hat{P}_i^*).$$

The trial value for which $\chi_2^2 - \chi_1^{2*}$ equals 5.412 is the lower limit on the safe dose for the assay considered. With k dose levels, there will be k sets of assay data and k lower limits, the maximum of these limits being the one selected. (The likelihood ratio may be taken as an alternative criterion for setting limits; *see 10.*)

REFERENCES

(1) Miller, G. B.: The role of certain chemical and physical agents in the causation of cancers. Hearings before the Committee on Interstate and Foreign Commerce, 86th Congress, 2d Session, on H.R. 7624 and S. 2197, January 26, 1960. Washington, U.S. Govt. Print. Office, 1960, pp. 45–60.

(2) Bliss, C. I., and Cattell, M.: Biological assay. Ann. Rev. Physiol. 5: 479–539, 1943.

(3) Finney, D. J.: Statistical Methods in Biological Assay. New York, Hafner Publishing Co., 1952.

(4) Burn, J. H., Finney, D. J., and Goodwin, L. G.: Biological Standardization, 2d ed. London, Oxford University Press, 1950.

(5) Bryan, W. R.: Quantitative biological experimentation in the virus and cancer fields. J. Nat. Cancer Inst. 22: 129–159, 1959.

(6) Federal Works Agency, Work Projects Administration for the City of New York: Tables of probability functions 2: 1942.

(7) Ordnance Corps Pamphlet ORDP 20-1: Tables of the cumulative binomial probabilities. Washington, Ordnance Corps, September, 1952.

(8) Abbott, W. S.: A method of computing the effectiveness of an insecticide. J. Econ. Ent. 18: 265–267, 1925.

(9) Counticld, J., and Mantel, N.: Some new aspects of the application of maximum likelihood to the calculation of the dosage response curve. J. Am. Stat. Assoc. 45: 181–210, 1950.

(10) Mantel, N., and Patwary, K. M.: Interval estimation of single parametric functions. Proc. 32d Session of the International Statistical Institute, Tokyo, Japan, 1960. In press.

(11) Bryan, W. R., and Shimkin, M. B.: Quantitative analysis of dose-response data obtained with three carcinogenic hydrocarbons in strain C3H male mice. J. Nat. Cancer Inst. 3: 503–531, 1943.

(12) Lea, D. E.: Actions of Radiations on Living Cells. Cambridge, Cambridge Univ. Press, 1946

(13) Meynell, G. G.: Inherently low precision of infectivity titrations using a quantal response. Biometrics 13: 149–163, 1957.

OGY AND APPLIED PHARMACOLOGY 20, 419–438 (1971)

Food and Drug Administration Advisory Committee on Protocols for Safety Evaluation: Panel on Carcinogenesis Report on Cancer Testing in the Safety Evaluation of Food Additives and Pesticides[1, 2]

August, 1970

Received April 26, 1971

INTRODUCTION

report is concerned with an evaluation of the present status of testing for carcino- action of food additives and other chemicals, which come into contact with man nally through his diet. Although carcinogenicity is only one of the many mani- ons of toxicity, it has somehow assumed a role of considerable, possibly magni- portance. It is the only manifestation of chemical intoxication that is dealt with by legislation. For practical purposes, a carcinogen is a substance that when admini- d by an appropriate route, causes an increased incidence of malignant tumors in mental animals as compared with a control series of untreated animals. The se to a carcinogenic substance may consist of (1) an increased incidence of tumors pe seen commonly in the control animals, (2) an occurrence of a type of tumor not t all in the control animals, or (3) a combination of the occurrence of a different of tumor and an increased incidence of one or several types of tumors seen in the

views expressed in this report are those of the Advisory Committee and are not to be regarded l statements of the U.S. Food and Drug Administration.

DA Advisory Committee on Protocols for Safety Evaluation.

ers of the Committee: Norton Nelson, Ph.D., Director of the Institute of Environmental ine, New York University Medical Center, New York, New York 10016, *Chairman*; Julius n, M.D., Professor and Head of Pharmacology Department, Jefferson Medical College, hia, Pennsylvania, 19107; Jerome Cornfield, B.S., Research Professor of Biostatistics, Pittsburgh Graduate School of Public Health, FASEB Building 9650 Rockville Pike, Bethesda, 3014. * Leo Friedman, Ph.D., Professor of Nutrition and Food Safety, Massachusetts e of Technology, Cambridge, Massachusetts 02139 (Present address: Director, Division of ology and Toxicology, Bureau of Science, Food and Drug Administration, Washington, 20204). * Robert E. Gosselin, M.D., Ph.D., Professor and Chairman of Department of ology, Dartmouth Medical School, Hanover, New Hampshire 03755. Calvin M. Kunin, Associate Professor of Preventive Medicine and Medicine, University of Virginia School of ne, Charlottesville, Virginia 22901; * Ted A. Loomis, M.D., Ph.D., Professor of Pharmacology, ity of Washington School of Medicine, Seattle, Washington 90105; Philippe Shubik, Ph.D., R.Phil., Director, Eppley Institute for Research in Cancer, University of Nebraska College ine, Omaha, Nebraska 68105; James L. Whittenberger, M.D., Professor of Physiology, rd University School of Public Health, Boston, Massachusetts, 02115; James G. Wilson, M.D., Professor of Research Pediatrics and Anatomy, Children's Hospital Research Foundation, nnati, Ohio 45229.

ive Secretary: James R. Cribbett, Committee Management Officer, Bureau of Science, Food and Administration, Washington, D.C. 20204.

bers of the Panel on Carcinogenesis: Dr. Philippe Shubik, *Chairman*; * Dr. Leo Friedman, Dr. on Nelson, Dr. James L. Whittenberger, Jerome Cornfield.

Consultant to the Panel on Carcinogenesis: William Lijinsky, Ph.D., Eppley Institute for Research in cer, University of Nebraska College of Medicine, Omaha, Nebraska 68105.

Membership terminated June 30, 1969.

controls. In some experiments the only manifestation of an effect consists of (4) an increased occurrence of tumors in the treated animals than in the controls, the incidence being the same in both. In yet another variation the only effect seen may consist of (5) an increase in the number of tumors per animal, the number of tumor bearing animals being the same. Any one of these effects may suggest the presence of a carcinogenic hazard; occurrence of benign tumors only would not classify an agent as a carcinogen.

It is apparent that all effects which in one way or another influence the occurrence of neoplasms cannot be precisely the same. The different situations cited have been examined in detail by experimentalists interested in mechanisms of action of carcinogens; however, from the standpoint of the toxicologist many different effects may be lumped together in an effort to simplify a practical task.

In a brief consideration of the basic research aspects of this problem in terms of practical uses made of the determination of "carcinogenic action" in animals, a set of problems will be posed and discussed:

1. Are Tumors Ever Induced That Do Not Occur in Control Animals?

The primary problem faced in answering this question is that relatively small numbers of animals must, of necessity, be used. When a rare tumor of man is considered, say a tumor with an incidence of less than 1% of all tumors occurring in man, it must be remembered that it is only by examination of a very large population that such tumors are discovered. There are on record various examples in which an experimentalist has concluded that he has induced a new lesion for the first time only to discover the lesion occurring in his untreated controls at a later date. Since many of the carcinogens used in the experimental laboratory are common environmental contaminants, albeit at low concentrations, it is not surprising that continued observation of "untreated" animals often yields examples of all the induced tumors seen in test animals.

It is thus ill-advised to conclude that a tumor seen in treated animals *never* occurs in the controls. The implication of this conclusion is that the occurrence of one or very few lesions, however rare, cannot be used with confidence to classify a compound as a carcinogen.

2. Are Benign Tumors of Significance to the Toxicologist?

In the first instance benign tumors may cause death in man and animals without ever undergoing malignant transformation. The induction of a benign tumor is, itself, therefore, an indication of a serious adverse reaction. There can be no doubt from a survey of experimental studies that benign neoplasms are often precursors of malignancies.

Under these circumstances, it would be wise to take serious note of the occurrence of benign neoplasms in experimental studies, although this alone is not sufficient for a conclusion of carcinogenesis. The occurrence of metastases provides an unequivocal demonstration of malignancy. There are, however, many tumors induced experimentally which are invasive and are classified as malignant, that metastasize only rarely during the average experiment. The absence of metastasis, in the view of most pathologists, does not rule out the diagnosis of malignancy.

Transplanatation has been used by some investigators as additional proof of malignancy. This method is employed infrequently and has certain drawbacks. There are, for example, some tumors such as the mammary fibroadenoma of the rat that are benign in all respects and yet may be transplanted readily.

at Is the Significance of Viral Oncology to Tests for Chemical Carcinogens?

time goes on, more and more tumors of laboratory animals have been found to be ated with identified viral agents. Most notable are the lymphomas of the that have been shown in some instances to be causally related to a specific

tain experimental models have demonstrated that lymphomas induced in mice by g radiations (Lieberman and Kaplan, 1959) or by a chemical carcinogen (Toth, may sometimes be transmitted by a virus. Other studies have demonstrated that tumors known to be transmitted by a virus in mice may be made to occur sooner application of certain chemical carcinogens (Kirschbaum *et al.*, 1940); these same omas, however, are not influenced by all carcinogens (Toth and Shubik, 1967). have been many studies to determine the effect of combinations of chemical ogens and viruses which are not necessarily carcinogenic. Some of these studies and Wisely, 1963; Duran-Reynals, 1952) indicate that it is likely that inter- between viruses that are not usually considered oncogenic, and known chemical gens may be important in the occurrence of neoplasms. The problem is compli- by the fact that, in recent years, various common viruses (e.g., adeno 12) have been to induce tumors when injected into infant rodents. These viruses have been ded only with nonneoplastic conditions in man. In addition, a ubiquitous virus, n as the polyoma virus, has been found to infect many mice although it is not ently associated with the occurrence of neoplastic disease in these animals in the ate. This polyoma virus can, however, give rise to many varied neoplasms when ed into newborn mice and hamsters.

are, therefore, faced with serious problems insofar as testing and interpretation ncologic purposes are concerned. Is there any value in a test of a chemical ogen in the Akr mouse? This mouse has a high incidence of lymphoma known to nsmitted by an identified virus. If an enhancement of this lymphoma occurred it be difficult to extrapolate such results to man since, up to this time, no human ancy is known to be transmitted by a virus. On the other hand, many scientists that leukemia may eventually prove to be a virus disease in man, and in the in- of African Lymphoma (Burkitt tumor) evidence continues to accumulate to this

e suggestive evidence that viruses and chemical carcinogens may interact in a s of ways indicates that tests of chemical carcinogens with viruses administered mitantly or demonstrated to be in the animal model may be of considerable tance. However, for the time being it would seem essential that until further basic s elucidate the significance of these phenomena, such studies cannot be recom- ed routinely.

the present state of our knowledge it is important that the toxicologist be familiar the status of viral oncology and that the viral background of animals used for ing for carcinogens be established. Although standardization of animal colonies be difficult, investigators performing carcinogenesis bioassay should know as has can be determined about the viral, bacterial, and parasitic contamination of the als. Advice on studies directed toward the development of defined laboratory als for bioassay can be provided by the Laboratory Animal Resources of the onal Academy of Sciences—National Research Council.

4. Is It Important to Take the Genetic Background of the Test Animals into Acco

A common criticism of animal studies that reveal a carcinogen is that the study undertaken in animals with a high tumor incidence and that the study thus applies to sensitive animals. Conversely others allege that a study was not satisfactory b the particular animals used were notoriously insensitive to carcinogenic a Cancer researchers are extremely fortunate compared to many other biologists in available to them many inbred strains of mice with known incidences of neop disease. In many instances there is also a considerable background of informa the sensitivity of these strains to the effects of known carcinogens. It is thus possi test unknown compounds that resemble known carcinogens chemically in rela small groups of animals and to derive meaningful results.

Shimkin *et al.* (1966) has recently suggested that investigation of compounds in re A mice and gauging the response by the incidence of lung adenomas might consti sensitive and accurate screening test for carcinogens. This contention requires t experimental study before it can be recommended generally. However, for cer specific situations there is no doubt that many of the special strains available a some instances, specially bred hybrids, provide a much more homogeneous popu to study and obviate many of the problems in evaluation of effects that are s randomly bred animals.

There are, nevertheless, many investigators who prefer a randomly bred anim the basis that the genetic heterogeneity of such animals is more akin to the situati man. There is, of course, much to be said for this viewpoint provided it is recog that often many more animals are required under these conditions

Only in a very limited number of cases have the particular tumors occurri inbred animals (AKr lymphomas, C3H mammary tumors) been shown to be assoc with a viral transmission. In the remaining instances the precise mode of origin tumors remains mysterious. The genetic influence has been established and fu studies are needed to uncover the mechanisms concerned. Under these conditions difficult to construct experiments that have clearly demonstrable relevance to p ticular situations in man. Special strains can be used for special purposes and pos incorporated into screening tests.

5. What Are Cocarcinogens and Promoting Agents?

The term cocarcinogen was coined by Shear (1938) to describe an enhancement the local carcinogenic effect of a weak solution of benzo[*a*]pyrene on mouse skin noncarcinogenic solution of a fraction of creosote. Various other noncarcinogens already been shown to have a similar action. Notable among these was the un vegetable mixture, croton oil. When mixed with a solution of benzo[*a*]pyrene below the concentration needed to induce skin tumors in mice, croton oil, noncarc genic alone, augmented the action of the carcinogen and gave rise to many tum Subsequently it was found that a single noncarcinogenic dose of the carcinogen c be followed by repeated applications of croton oil and that many tumors would re (Terracini *et al.*, 1960). Similar experiments in rabbits showed that carcinogenesis the skin could be begun by applications of coal tar or methylcholanthrene and c pleted by wound healing or applications of turpentine. This two-stage induction of sk

...was the basis for a hypothesis describing the phases as "*initiation*" and "*promo-*
...The promoting agent croton oil has proved to be the most potent so far dis-
...it has been fractionated, and interesting substances with this action isolated.
...bstances with less activity on the skin have been found (Lijinsky and Saffiotti,
...affiotti and Shubik, 1956). The two stage mechanism has not been convincingly
...rated *in carcinogenesis* in sites other than the skin. There are many situations
...a two-stage mechanism appears as an attractive explanation. We are still at a
...ere much more study is needed before many of these concepts, valid for the
...be said to have general application. From a practical standpoint it is apparent
...on oil or substances of similar activity would pose a considerable hazard to the
...Although such interactions may pose an important public health problem, our
...state of knowledge does not make it practical to recommend routine tests for
...nogenicity or promoting action; if circumstances suggest the existence of such
...testing should take this into account. This area needs intensive study.

...*ources and the Total Environmental Load of Carcinogens*

...es on record that are concerned with the effects of mixtures of chemical
...gens have, in the main, reported the occurrence of inhibition of tumor
...on. Lacassagne *et al.* (1945) reported that polycyclic and heterocyclic hydro-
...s administered together resulted in inhibition; this was thought to be due to
...tive inhibition for similar receptor sites. Subsequently these studies were re-
...and extended with essentially similar findings. Richardson *et al.* (1952) later
...trated that administration of methylcholanthrene could inhibit the induction
...omas in rats by *p*-dimethylaminoazobenzene. Many additional studies by
...investigators (Gelboin, 1967) demonstrated that this was due to enzyme
...on. Since then many instances of effects of enzyme induction on the action of
...gens have been reported. Many students of carcinogenesis have drawn attention
...lack of information on the effects of low doses of combinations of different
...gens in animals. It has been said that the practical conditions of human exposure
...must involve the presence of many carcinogens of different types has not been
...d in animal studies.

...e are many known carcinogens that are chemicals, physical agents and viruses
...in our environment. Some of these agents are known hazards to man; others are
...own to be active in animals. It has become apparent to some researchers that it is
...if not impossible, to assess the hazard from any single carcinogen out of the
...of the total environment.

...polycyclic aromatic hydrocarbon carcinogens illustrate the point. There are
...polycyclic aromatic hydrocarbon carcinogens present in coal tar, certain petro-
...oducts, tobacco smokes and tars, polluted urban atmospheres, charcoal-broiled
...oked foods and other products. Obviously the discovery of an additional source
...ilar quantities of these compounds could be said to contribute to a general
...However, the assessment of the hazard must take into account the total
...e. In the instance of hazard from ionizing radiation a relatively arbitrary con-
...estimating "hazard above background" has long been used. Eventually a similar
...ach to the assessment of hazard from chemical carcinogens would seem indicated.
...an approach is even more difficult in this instance because of the difficulties of

quantifying the levels of the many different carcinogens present in the environm... the present time the committee on Quantitative Carcinogenesis of the Interna... Union Against Cancer (UICC) has set itself the task of providing a tabulation ... known quantities of chemical carcinogens in the environment and listing anal... methods. This task is formidable but underlines the need for more information be... philosophical basis for the evaluation of hazard from carcinogens can be establis... a rigorous manner.

7. Are There Special Features Associated with Hormonal Carcinogenesis?

The refusal to establish a legal tolerance for the herbicide aminotriazole raised ... questions that have not been answered to the satisfaction of many toxic... Aminotriazole is one of a series of thyroid-stimulating agents that give rise to ... hyperplasia, which in turn can lead to the occurrence of thyroid adenomas and ... nomas. Additionally such effects are intimately bound to the total hormonal ... the animal and may be accelerated or inhibited in an indirect manner. In the insta... stilbestrol, a substance with hormonal activity of great interest to the toxic... interested in food additives, the situation is similar. Prolonged administration ... hyperplasia in the first instance, leading to the occurrence of benign tumors and ... carcinoma. However, in the instance both of certain of the thyroid-stimulating ... mones and of certain estrogens the induction of hepatomas, not apparently ass... directly with hormonal effects, can occur.

In the main, however, hormonal carcinogens are distinguished from other ... nogens in that, the final result is preceded by an exaggerated physiological effect.

The problem of assessment of hazard from hormonal carcinogens is, therefore, ... less complex than any other problem in chemical carcinogenesis. The hazard ... exposure to hormones must be evaluated both in the context of the hormonal ... and as chemical compounds. When the tumors induced are clearly related to phy... logical effects it may be possible to assess dose response relationships and other ... tions of exposure on a reasoned basis. However, in the instance of many hor... other effects of a chronic nature may occur which have not, so far, been associated ... the known physiological role of these agents.

8. The Role of Pathological Reactions to Inert Materials

Physical carcinogenesis in the subcutaneous site has been extensively studied ... well known that a variety of chemically inert materials may be introduced ... cutaneously in rodents and will give rise to sarcomas after a relatively long ... period. Most such studies demonstrate that the ability to induce tumors is related to ... physical nature of the material (hardness, perforations, configuration, etc.), and, ... the mechanism remains unknown.

In certain studies with food additives the occurrence of bladder calculi subseq... associated with tumors of the bladder has been the subject of much discussion ... case of the compound polyoxyethylene monostearate (Myrj 45) bladder calculi ... reported to occur at very high oral dose levels; some of these animals developed bla... papillomas. It has been widely assumed that the two events were related although ... has not been proved. It has been shown, however, that glass beads inserted into ... urinary bladder could give rise to papillomas in rodents. It can, therefore, be reasona...

...med that in the My... ...rrence of the calcu... ...n a secondary ma... ...r conditions where ...ied and conclusion...

DISCUSS...

One of the primary ...sideration of recon... ...e most of the wr... ...ances have been m... ...e been performed ... ...propriately incorpo... ...ssions of the com... ...ndustry and the ac...

...the Desirability of ...nimals in Routine

...here is agreement ... The use of SPF ...imals, as free from ... ...based SPF anima... ...e conditions the S... ...eptible to infectio... ...there is considera... ...ed with conventio... ...ntiate the view that ...'s for chronic toxi...

... The Bladder Impl...

...Several chemicals ... ...nificant increase i... ...ntrol animals imp... ...ich increase the i... ...mic. This assumpt... ...monstrate safety ... ...ed both with thi... ...ble secure interp... ...ne by this assay ...ould be obtained ... ...ined only with ...cinogenicity.

...Since this report w... ...mission of the UIC... ...ver, Geneva 1969, ...abeth Weisburger, ...ss, 1967.

...d that in the Myrj 45 finding the induction of the tumors was secondary to the ...nce of the calculi and, therefore, that the compound manifests tumorigenicity ... secondary manner. It would seem logical that tumors arising experimentally ...nditions where physical factors might play a significant role should be carefully ... and conclusions based on common sense evaluations.

DISCUSSION OF UPDATING TESTING PROCEDURES

...of the primary tasks to which the committee has addressed itself has been a ...ration of recommended test procedures for carcinogenesis. It is over a decade ...most of the written recommendations were made.[3] During that time many ...s have been made in biological knowledge and techniques; many experiments ...een performed in carcinogenesis. It was felt that this new knowledge must be ...riately incorporated in test procedures. The following proposals emerge from ...ons of the committee and its consultants and after a meeting with representatives ...stry and the academic community.

...e Desirability of Introducing Specific Pathogen-Free (SPF) Animals or Germfree ...mals in Routine Tests

...e is agreement that germfree animals are not required or desirable for routine ... The use of SPF animals is also *not recommended* as a requirement. Healthy ...s, as free from infection as possible, are recommended. Some investigators have ...ed SPF animals and used them in animal quarters with standard animals. Under ...nditions the SPF animals were, in fact, less satisfactory since they were more ...ible to infection than standard animals.

...re is considerable dispute as to the length of time that SPF animals survive com... with conventional animals. There does not seem to be adequate evidence to sub...re the view that the introduction of SPF animals might lengthen the period avail... for chronic toxicity testing.

...e Bladder Implantation Method of Assay for Carcinogenic Activity

...ral chemicals which are carcinogenic by other tests have been found to produce a ...ant increase in the incidence of bladder carcinomas above the incidence found in ...d animals implanted with the pure vehicle. It is assumed that other chemicals ... increase the incidence of bladder tumors with this technique are also carcino... This assumption may be correct. However, at this time it is not adequate to ...strate safety of a chemical with this as the only test. The number of chemicals ...both with this technique and with more conventional methods is inadequate to ...secure interpretation at present. However, a chemical which is clearly carcino...by this assay must be regarded with considerable suspicion, and further data ...be obtained with more conventional assays. More important, negative data ...ed only with this method cannot be accepted as adequate evidence of lack of ...genicity.

...e this report was prepared, "A Report of the Panel on Carcinogenicity of the Cancer Research ...sion of the UICC," *Carcinogenicity Testing* (I. Berenblum, ed.), International Union Against ..., Geneva 1969, has appeared. See also, "Tests for Chemical Carcinogens," by John H. and ...th Weisburger, in: *Methods in Cancer Research* (H. Busch, ed.), Vol. I, pp. 307–398. Academic ... 1967.

3. The Use of the Dog in Tests for Carcinogenicity

There is general agreement that the dog is not a practical test animal because size and relatively long life span. It is felt, however, that all substances related aromatic amine carcinogens should be tested in this species. It is pointed out drug chlornaphazine (*N,N*-bis-2-chlorethyl-2-naphythlamine) has been shown potent human bladder carcinogen. Had this drug been tested properly in the to clinical trial, this occurrence could have been prevented. In the instance of comp related to the aromatic amine-bladder carcinogens, the dog test lasting at least is needed under present conditions. It seems likely that the hamster may prove satisfactory substitute.

4. The Use of Species So Far Not Commonly Used in Such Tests

There is general recognition that new species in addition to the common rodents are needed for carcinogenicity testing. Increasing the number of spec would greatly increase the margin of confidence with which safety can be pr Up to this time there have been no species that can be recommended with co A variety of species show promise, but in all instances further study is needed they can be recommended for routine test purposes. The most discussed ad species are primates. Recent studies show certain primates to be highly susce some carcinogens; however, much more work is needed to justify the additional and time that would be necessitated by the introduction of these animals in rout procedures. Among the other species that have been suggested are fish, partic rainbow trout, and various avian species such as chickens, parakeets, a Japanese quail. At present there are no specific recommendations that any of species should be considered for routine testing.

5. General Environmental Conditions

Although Committee attention has been directed specifically to air contamin animal quarters in studies of chemical carcinogenicity, it is evident on reflection is only one part of the environmental system that can be inadvertently conta with the chemical under test or other carcinogens in the environment. Thus, a under test may enter the general laboratory atmosphere from faulty technique preparation or by secondary aerosolization of animal wastes. Other examples are tamination of diet, air, or contact surfaces by chemicals used for insect or control, or from preceding tests.

Although it would be inappropriate in a document of this kind to attempt to specific requirements for control of environmental contamination by potential nogens, the Committee emphasizes the extreme importance of good animal care ties in any long-term study of low grade toxicity or carcinogenesis. Details of st caging, and environmental control may vary from one facility to another, but the tial ingredient of good animal care is the personal component—the well-instr highly motivated caretaker staff who are genuinely concerned about the well-be the animals.

Minimal standards for animal facilities and care are now specified by law. Det guidelines, including special requirements when toxic, infectious, or radioactive

are used, are available in published documents such as the Guide for Laboratory Facilities and Care (NRC 1968).

Use of a Standard Diet in Experiments for Carcinogenesis

ome safety evaluation studies, including carcinogenesis tests, have been based on diets of practical ingredients supplied commercially by laboratory feed manu- or, on occasion, mixed from practical ingredients in the investigator labora- Concern with this basal diet has been limited to (1) meeting the requirements for growth, reproduction and longevity and (2) maintaining a constant composi- oughout the experiment. Usually the constancy of composition is assured only of the proximate analysis, vitamin content, and occasionally, amino acid con- In 1953, Tannenbaum and Silverstone wrote: "... natural foods contain a r of constituents which have been given little attention in nutrition and cancer h because they are apparently not dietary essentials. In addition, there must be yet undetected. Perhaps among these unregarded substances are some with genic activity; and others that potentiate or oppose the action of carcino- "

e then, many examples to prove the truth of these words have accumulated. there is an awareness of the presence in foodstuffs of many nonnutrient com- with interesting biological activities, including carcinogenic activity. It is that in the evaluation of carcinogenic potency, as in any bioassay, it is essential d "contamination" from environmental sources which may produce erroneous

ccurrence of nonnutrient components in foodstuffs and their influence on bio- response have recently been reviewed by Friedman *et al.* (1966), Schoental (1965), tering (1966)

clear that a complex diet of practical ingredients is difficult to maintain constant position over a long period of time with regard to all ingredients. Even with the stringent quality control, one must expect variation in the same crop grown on the oil from one growing season to another. How much more difficult it is to maintain stancy of composition with regard to ingredients gathered from many sources any growing periods, different harvesting situations, different storage and pro- practices, different pesticide residues and residue levels, etc. Experience tells us contamination" of the type that would significantly influence carcinogenic studies serious (Newberne, 196 ; Saracini *et al.*, 1967). Often the interpretation of studies is extremely diff because the incidence of malignant tumors in con- mals approaches that ob ed in the test groups, forcing a conclusion that the oup did not differ significantly from the control. Such a conclusion is consistent the presence of low level carcinogenic activity in the test group or in both groups. rmore, what can be the significance of the incidence of "spontaneous" tumors in ible strains when one is not certain about the presence of carcinogenic con- nts in the diet on which the animals have been maintained?

essential nutrient requirements of many laboratory animals, particularly that of and the mouse, have been identified to the point where it is possible to maintain on chemically defined diets. Somewhat simpler and more economically feasible ade up of purified ingredients, such as purified casein, sucrose, starch, salt mixes,

and synthetic vitamins have been used by many investigators. Halver *et al* (19[illegible] used such diets in studies of hepatoma over a period of several years in rainbow Wogan and Newberne (1967, 1968) have recently successfully completed an IS study in rats on the carcinogenic potency of aflatoxin B_1 using a diet very simila used by Allison *et al.* (1950, 1954). The feasibility of using highly purifie synthetic diets mixed in the investigator laboratory under his continuous cont been demonstrated. The advantages of such a basal diet in carcinogenesis stud the point of view of maintaining constancy of composition with regard to a ponents, and particularly with respect to elimination to the highest practical de all potential dietary carcinogens and other factors that may influence carcino seem to be sufficiently compelling to recommend its adoption in all carcinog evaluation.

The increase in cost of feed, although appreciable (from approximately 10 pound to approximately 50-60 cents a pound) represents only a minor fraction total cost of a carcinogenesis study. It must be emphasized that as much care be exercised in preparation of basal diets as in the preparation of experimental containing test substances.

7. Pathological Study

It is of the utmost importance that a complete and accurate pathological exami be conducted on all animals used in carcinogenesis experiments. A trained path should be concerned with both the gross and microscopic diagnosis of all neop found. The pathologist should have a part in designing and conducting the experi The autopsy should be done under his general supervision. It is essential that the p logist be experienced in the diagnosis of animal neoplasms; knowledge of human t alone is not a sufficient background for expertise in this field. It is desirable that gross and microscopic examinations be conducted without knowledge of the treat of specific animals.

Many pathologists employ a check list to ensure that a complete autopsy has employed, and it is felt that this is a good procedure.

8. Rapid Tests

There is no rapid test that the Committee recommends for routine screeni carcinogenic chemicals. The most promising procedure is tissue culture. This meth used routinely for the evaluation of oncogenic viruses, where "transformation" of cultures by viruses is used as proof of oncogenicity. In its inception, "transforma was a phenomenon characterized by the development of unlimited growth potent a culture on continuous transfer, and by loss of contact inhibition as could be see the architectural pattern of growth. The proof of the changed characteristics of culture was the induction of sarcomas on reinoculation into an animal. Since that only the manifestations of lack of contact inhibition have been used as the indicator in the instance of certain experiments conducted with well established lines of cultures experienced virologists, valuable comparative data on the characteristics of diffe viruses and their mutants may be obtained. However, these findings cannot yet be as an absolute diagnosis of carcinogenicity of an unknown in the absence of all stan

...cedure including reinoculation. Even when this is done, efforts must be made to ...e that the possibility of "spontaneous transformation" has been excluded.

...spite of these strictures more effort should be made to follow up recent investiga... ...by Berwald and Sachs (1963), Heidelberger *et al.* (1968), and others on the use of ...cultures. It must be cautioned that many carcinogens are metabolized to the active ...ound *in vivo*. The activity of such compounds might be missed in an *in vitro* ...n

...e of Newborn Animals*

...e newborn mouse and hamster have been used widely for the study of both ...genic chemicals and viruses. In some studies, it was found that the newborn ...was more sensitive to the rapid induction of lymphomas and lung adenomas than ...dult. This was correlated with a longer persistence of the unchanged molecule of ...rcinogen in the newborn than in the adult. In the case of viral agents the lower level ...mune response in the newborn is considered of major importance, and the same ... may also play a role in the instance of the chemicals.

...ain compounds such as fluorenylacetamide that are metabolized to a proximate ...gen may be less active in the newborn than in the adult. Therefore, the newborn ...e has limited applicability as an overall test animal for carcinogenicity. In some ...ses, it may be convenient to use newborn animals since relatively small quan... ...of materials may be administered only once or a limited number of times. The ...ls can then be maintained without further treatment and tumors be observed ...only a few months in the case of very active materials. However, this test does not ...n the period of total observation required for a negative result. Indeed there are ...ses on record in which tumors have occurred very late in life after administration ...rcinogens at birth. These include the induction of (a) hepatomas by DMN in mice ... *et al.*, 1964) and by urethan in mice (Della Porta *et al.*, 1967), (b) subcutaneous ...mas by DMBA in rats (Toth and Shubik, 1963), (c) kidney and liver tumors by ...in rats (Hirono *et al.*, 1968), and (d) neurogenic tumors (brain and peripheral ...by urethan in rats (Vesselinovitch and Mihailovich, 1968). It has been suggested ...newborn animals might be treated at birth and then followed for only one year. ...might be a satisfactory screening procedure under limited conditions but cannot ...commended as a routine test procedure.

...e Combination of the Chronic Toxicity and Reproduction Study

...e evidence is striking that sometimes administration of a single dose to 1-day-old ...s or of several doses beginning on the first day of life and at intervals up to time ...aning, and observation up to one year, constitutes a very sensitive assay system for ...tection of carcinogenic potential. The knowledge, for example, from the work of ...ovitch *et al.* (1966), and Druckrey *et al.* (1967), that tumors may be produced late ...life of animals after administration of a carcinogen during the pregnancy of the ...r indicates that the embryo and fetus may sometimes be particularly vulnerable ...s for carcinogenic substances.

...the context of a safety evaluation protocol in which subtle chronic effects are ...mary concern and difficult to demonstrate, and in which the purpose is to assure ...mplete a screening of as many potential deleterious effects as possible, this

information on the susceptibility of perinatal animals suggests some modifica in the protocols of chronic toxicity studies.

The chronic toxicity studies may be carried out independently or coordinated the reproduction studies, at the convenience of the laboratory. At present, studies, including carcinogenesis, begin with weanling animals derived from a stock colony, maintained on adequate practical diets. These animals are then tained at desired dose levels for a "lifetime," usually a 2-year period for rats. They received *no perinatal exposure*. Since one of the important purposes of the toxicity tests is the detection of carcinogenic potential, it would seem desirable to the exposure as early in life, i.e., as close to conception as possible. Because one reasons for a reproduction study is the detection of effects in the offspring, from exp of the parents before and during conception and gestation, it would seem probab combine the first stages of the reproduction study with the chronic toxicity test the initial animals started in the experiment at weaning would be bred and the offsp either used solely for the 2-year chronic toxicity test or continued for the remain the 2-year period to serve as chronic toxicity test animals as well as the F_1 gener after producing the second generation of the reproduction study. This would all advantages of perinatal exposure for carcinogenesis studies, and the chronic toxicity carried out in this manner may prove to be an adequate test for carcinogenic hazard

11. Statistical Aspects of Safety Evaluation

There are two questions that may be asked about an agent under test:

a. For a given set of experimental conditions (e.g., total dose, fractionation, route administration), is there any elevation above control levels in the incidence of cancer i.e., is the agent carcinogenic?

b. If yes, can a dosage be found below which there is no elevation, all other experimental conditions remaining the same; i.e., even though carcinogenic, can a safe be found?

Although we are concerned with carcinogenicity, it is clear that the formal argume largely applies to any other adverse effect of interest as well. Extrapolation of long chronic effects from animal to man introduces more uncertainty that extrapolation acute effects.

We consider statistical aspects of each question in turn.

Is the Agent Carcinogenic?

Although a positive answer to the question as posed can be given in some partic instances, no unqualified negative answer is ever possible. Thus, if none out of n animals exposed to the agent develop cancer, and n is a large number, the elevati any, is certainly small, but no matter how large the value of n, it is not logically po to conclude on the basis of such evidence that *no* elevation has occurred. The might in fact induce one cancer in every $2n$ animals, and the experiment, being too to detect it, would be incapable of providing cogent evidence for the absence of elevation whatsoever. This is true no matter how large is the value of n.

A negative experiment cannot therefore be regarded as having shown that no ~~ti~~ tion exists. It can, however, supply an upper limit to the possible carcinogen Clearly, the larger the value of n, the closer this upper limit will be to zero, but there

...ate n that will set it eq
...mpute this upper lim
...hich cancer will be ind
...his proportion may ha
...s one animal will dev
...at all n test animals wi
...lue of P is the probab
...his is easily obtained
... $1 - 0.01^{1/n}$. The
...known value of P. Va
...below.

UPPER CONFID
WHICH A

n
10
100
1000

These upper limits a
...nd using only 90% co
...s 23 cancers per 1000
...uman population for
...more than 2 tumor
...illion with confidenc
...more than three milli
The situation is not
...ble. We show below
$P_1 - P_c$ where the sur
proportion of animal
...qual for treated and
P_1 for the case in whic

UPPER 95% C
EXPERIMENTS
SAME INC

n
10
25
100
500
1000

* The approximation

n that will set it equal to zero. An elementary probability argument shows how to pute this upper limit in the simplest case. Denote the proportion of animals in h cancer will be induced by the agent under the given experimental conditions by P. proportion may have any value from zero to unity inclusive. The probability that ne animal will develop cancer is P and that it will not is $(1 - P)$. The probability all n test animals will fail to develop cancer is then $(1 - P)^n$. We then ask: For what e of P is the probability of the observed outcome of zero out of n small, say 0.01? s is easily obtained by setting $(1 - P)^n$ equal to 0.01 and solving for P to obtain $= 1 - 0.01^{1/n}$. The interval 0 to P_{max} then provides a 99% confidence limit on the known value of P. Values of P_{max} for various n and confidence coefficients are shown ow.

$n \log_{10}(1-P) = -2$

$1 - P = 0.01^{1/n}$

$P = 1 - 0.01^{1/n}$

UPPER CONFIDENCE LIMIT FOR POSSIBLE INCIDENCE FOR EXPERIMENTS IN WHICH ALL n ANIMALS TESTED ARE UNAFFECTED, BY n AND CONFIDENCE COEFFICIENT

n	90%	95%	99%	99.9%
	P_{max}			
10	0.21	0.26	0.37	0.50
100	0.023	0.030	0.045	0.067
1000	0.0023	0.0030	0.0046	0.0069

ese upper limits are uncomfortably large. Even with as many as 1000 test animals ing only 90% confidence limits, the upper limit yielded by a negative experiment cancers per 1000 test animals. No one would wish to introduce an agent into a population for which no more could be said than that it would probably produce re than 2 tumors per 1000. To reduce the upper limit of risk to 2 tumors per one n with confidence coefficient 0.999 would require a negative result in somewhat than three million test animals.

e situation is not improved when concurrent observations on controls are avail- We show below the approximate upper 95% confidence limit on the difference P_t where the subscripts stand for treatment and control, and P is, as before, the portion of animals developing cancer. This limit is shown for various n (assumed l for treated and controls with $2n$ equal to the total number tested) and values of r the case in which treated and control groups develop the same number of tumors.

UPPER 95% CONFIDENCE LIMIT FOR POSSIBLE ELEVATION INCIDENCE FOR EXPERIMENTS IN WHICH n TREATED AND n CONTROL ANIMALS YIELD THE SAME INCIDENCE, BY n AND CONTROL INCIDENCE (Cornfield, 1954)

n	0.01	0.10	0.25	0.40
10	0.10[a]	0.31[a]	0.45	0.51
25	0.065[a]	0.19	0.28	0.32
100	0.033[a]	0.098	0.14	0.16
500	0.015	0.044	0.063	0.072
1000	0.010	0.030	0.044	0.051

[a] The approximation used introduces some error in this range.

Thus, even with 1000 treated and 1000 control animals, each developing 10 tum... ... can be 95% sure only that the elevation in incidence rate does not exceed 1"... ... the agent has probably induced at most 1 additional tumor per 100 test animals. ... be observed that the practical difficulties in obtaining a completely unexposed ... are many. As has been noted in other parts of this report, the environment of the ex... mental animal (diet, water, and air) are often uncontrolled and may, at times, intro... low levels of the test substance itself or other carcinogens.

Considerable additional safety is of course assured by testing doses well above ... actual use level. But this leads to the question of a safe dose, consideration of w... seen to be an unavoidable consequence of the astronomically large number of a... required to reduce the upper limit of possible carcinogenicity to an acceptable le...

Can a Safe Dose Be Found?

The above calculations make clear that the only practicable basis for estimating a ... dose is by extrapolation downward from results obtained at some level well ab... actual use level. But this extrapolation introduces serious uncertainties, which m... recognized if rational methods of safety evaluation are to be developed. The basic p... lem is that extrapolation outside the range of observation must be based on pere... unverifiable assumption about the mathematical nature of the dose-response rel... ship near zero dosage. The variety of dose-response curves possible in the ... carcinogenesis model experiments further complicates the problem.

It might be thought that the basis for such an extrapolation could be provid... observations in the observable range. To show how far from being the case this act... is, we give below three different dose-response curves, mathematically defined ... dosage range of 256-fold. All three have the same TD_{50} and TD_{16}.[4] The first is a pr... curve, the second a logistic curve, and the third the so-called one-particle curve

EXPECTED PERCENT OF ANIMALS WITH TUMORS

Actual dose (TD_{50})	Probit curve (%)	Logistic curve (%)	One-particle curve (%)
16	98	96	100
8	93	92	99+
4	84	84	94
2	69	70	75
1	50	50	50
1/2	31	30	29
1/4	16	16	16
1/8	7	8	8
1/16	2	4	4

It will be noted that below the TD_{50} the three curves differ by little and that in any exp... ment of practicable size (fewer than several thousand animals), it would not be pos... to conclude from the actual observations which one of the three best described the d...

[4] TD_{50} = dose which results in tumors in 50% of the animals; TD_{16} = tumor dose for 16% of animals.

...s shown below, however, t... ...one hundred million do... ...arkedly.

EXTRAPOLATED VALUES ... CURVES DESCRIBIN...

$TD_{...}$	
$TD_{0.1}$	
$TD_{0.0001}$	
$TD_{0.000001}$	
TD_{1}	
$TD_{0.000001}$	

...the one in one-hundred ... ...virtually" safe dose is o... ...housandth using the logi... ...en with an experiment. ...virtually" safe dose is o... ...,000,000 depending on t...

The extreme unreliabil... ...as the basic source of t... ...cine (Cornfield *et al.*, ...activation time had so... ...arching to an extrapol... ...olecular basis as Boyle ...learly extrapolation fr... ...ties and imponder... ...prudent to place exces... ...the dose-response curves ...physical or chemical ba...

Nothing that has beer... What does seem clear is ... toxicologic dose-respon... ...n the question.

...tlines of a Broad Ge...

The impracticably lar... ...possible public health ... ...polation, set limits to ... testing program. Draw ... ...road outlines of a fea...

1. Testing should be ... maximum tumor incid... ...de above the actual ...

shown below, however, the $TD_{0.0001}$ (one in a million dose) and the $TD_{0.000001}$ (one in a hundred million dose) obtained by extrapolation of these three curves differ widely.

EXTRAPOLATED VALUES OF "SAFE" DOSES FOR THREE DIFFERENT DOSE-RESPONSE CURVES DESCRIBING OBSERVED RESPONSES IN THE 2% TO 50% RESPONSE RANGE EQUALLY WELL

	Probit curve	Logistic curve	One-particle curve
TD_1	0.040	0.022	0.014 4
$TD_{0.1}$	0.015 5	0.003 15	0 001 44
$TD_{0.0001}$	0.001 36	0.000 009 8	0.000 001 44
$TD_{0.000001}$	0.000 412	0.000 000 16	0.000 000 014
$\frac{TD_1}{TD_{0.000001}}$	100.0	100,000.0	1,000,000.0

one *in* one-hundred million dose, which Mantel and Bryan (1961) call the "safe dose is one-hundredth the TD_1 using the probit curve, one-hundred using the logistic, and one-one millionth using the one-particle curve. Thus, with an experimentally well-determined TD_1, or dose at which $P_{max} = 0.01$, the "safe" safe dose is obtained by using a safety factor which can vary from 100 to depending on the curve selected.

extreme unreliability of extrapolations outside the observable experimental range basic source of the failure of the early safety evaluation program for the Salk (Cornfield *et al.*, 1956), even though the observed curve connecting log titer and ation time had some theoretical physical chemical basis. Nor is the uncertainty ng to an extrapolated value surprising, since even a relationship with as firm a lar basis as Boyle's law breaks down at extremes of pressure and temperature. extrapolation from the observable range to a safe dose has many of the per- and imponderables of extrapolation from animal to man, and it would be nt to place excessive reliance on mathematical sleight of hand, particularly when se-response curves used are largely empirical descriptions, lacking any theoretical or chemical basis.

ing that has been said bears on whether a threshold dose does or does not exist. does seem clear is that more fundamental experimentation than that of the usual single dose-response investigation in intact animals is necessary to shed much light question.

es of a Broad General Strategy of Testing

impracticably large numbers of animals required to detect carcinogenic effects of public health importance, together with the uncertainties of downward extra- set limits to the amount of protection that can be achieved with a routine program. Drawing largely on the previous work of Mantel and Bryan (1961), the outlines of a feasible cancer testing program might appear to be as follows:

Testing should be done at doses and under experimental conditions likely to yield tumor incidence. This would mean the use of doses several orders of magni- above the actual use level. Although it is a simple exercise to calculate the number

of test animals required to achieve ideal protection, this number is likely to exceed bounds of practicability, and some practical compromise is called for. Thus, for special case in which the control incidence is zero, negative results on 100 test an at a high dose level would do no more than establish with 95% confidence that the dence does not exceed 3 in 100 at that dose level. With an incidence of 10 per animals in both controls and test animals one is 95% sure only that the elevation incidence with test does not exceed 10 per 100 test animals at that dose. Clearly no and fast rule can be laid down.

2. Positive evidence of carcinogenicity will be supplied when the lower confi limit on the elevation in incidence ($P_T - P_C$) is above zero. Expression of test confidence limits rather than as a test of significance is to be preferred, since even the lower confidence is below zero and no positive evidence exists, the upper limit be well above zero and this will serve as a constant reminder that failure to ob positive evidence of carcinogenicity is not the same as a positive demonstration of carcinogenicity.

3. For compounds judged carcinogenic at test levels, a virtually safe dose co principle, be estimated by downward extrapolation using some arbitrarily selec conservative dose-response curve. Mantel and Bryan (1961) suggest the use of a with a probit slope of 1 per 10-fold dilution on the grounds that all dose-resp curves with carcinogens are in fact steeper. But as they clearly recognize, use of rather than probits would be even more conservative and has an equal justifi (Mantel, 1963; Druckrey, 1967). A reasonable upper limit to carcinogenicity at use levels would be provided by use of a straight line connecting the origin and the corresponding to the TD_1. This is equivalent to saying that the expected proport tumors at some fraction, f, of the TD_1 is $f \times 0.01$ and would also follow from use one-particle curve. Any dose for which this expected proportion is less than some ar rarily low level, say 10^{-8}, would be judged safe. This would lead to few conflicts wi results of applying the Delaney clause. Unless the test level were at least 10^4 tim proposed use level, a finding of carcinogenicity at the test level would lead to so safe use level as to amount to a *finding of nonacceptability*. It is clear that the u taintics involved in extrapolating downward from test to use levels are not restric carcinogenesis and may apply to other toxic substances as well.

4. For agents not judged carcinogenic, such an estimation of safe dose would logical, but would be so low as virtually to exclude from use agents for which there no positive evidence of carcinogenicity. For such agents an alternative is a scaling ward from the test level by a more modest factor, say 100-fold. This is equivalent downward extrapolation using a probit curve with a slope of about 2 per 10-fold d should be clear that no absolute guarantee of the safety of such a use level could be It would be desirable to compute routinely the upper limit of *possible* elevation in dence at that use level by downward extrapolation, using the same conservative cedure recommended for agents found to be carcinogenic, i.e., by multiplying upper limit to excess elevation at that level by f, the fraction of the use to the test le Thus, if 100 controls and 100 animals tested at 100 times the use level, each yield tumor-bearing animals, the upper 95% confidence limit to the elevation at the level is 0.098 (Mantel, 1963, Table, p. 38) and at the use level 0.0098 (0.098 ÷ 100), about 1 per 1000.

...ommendations and Conclus

The committee recommends ...evaluated for carcinogenici ...information to be developed ...ancillary information ava... The decisions should be sub...

1. At least two species are ...a two rodent species be ...cinogens on record among ...eguards in consideration ... ...ly is considered inadequa... ...extend this requirement ... ...ative finding. No other ... ...gens to be recommended ...ould have to be evaluated

2. Randomly bred anim... ...ains should be reserved ...

3. Knowledge of the me... ...se of comparative metab... ...ay enable selection of te... ...riate, such information s...

4. It is essential that cor... ...et sufficient to rely on ... ...animals.

5. Positive controls ar... ...emically analogous to ... ...contamination of the test

6. For animal studies, ...uded which should b... ...ances such as additives ...mpound. It is essential ...ade to determine the l...

7. Testing should be ... ...aximum tumor incide...

8. Test numbers of an... ...that would not be acce... ...se of future testing to ... ...s, recognizing that t... ...ikely to exceed the be...

9. Although it is pos... ...udies involved in d... ...permissible levels that ...

10. The potentiatio... ...may be an important ... ...it impractical to rec... ...tion be conducted.

mendations and Conclusions

committee recommends that the FDA require that food additives and pesticides
uated for carcinogenicity. FDA scientists must exercise judgment as to the type
rmation to be developed. The requirement may vary with each case, depending on
illary information available.

decisions should be subject to periodic review by an independent advisory body.

least two species are recommended for all carcinogenicity tests. It is suggested
o rodent species be used. There are many examples of species variation to
gens on record among rodents. The use of more than one species adds additional
rds in consideration of extrapolation to man. The use of the dog for a two-year
is considered inadequate for the study of carcinogenicity, and it is felt unrealistic
end this requirement routinely to the seven or more years needed to establish a
ve finding. No other nonrodents have been sufficiently investigated with carci-
s to be recommended for routine tests at this time. Any new species or strains
have to be evaluated with a range of known carcinogens.

Randomly bred animals are recommended for routine use. The use of inbred
should be reserved for special tests.

Knowledge of the metabolic pathway of a test substance and particularly know-
of comparative metabolisms (quantitative as well as qualitative) in different species
enable selection of test species to be made on a logical basis, and, where appro-
, such information should be developed

It is essential that concurrent negative controls always be included in the tests; it is
fficient to rely on data of tumor incidence from prior studies with untreated
ols.

Positive controls are desirable when it is known that the test compound is
ally analogous to known carcinogens. Precautions should be taken to prevent
mination of the test animals by the known carcinogens.

For animal studies, a semisynthetic diet or a diet of known composition is recom-
ed which should be as free as possible from incidental discrete chemical sub-
es such as additives in individual ingredients, pesticides, and particularly the test
pound. It is essential that, if freedom from additives cannot be assured, analyses be
to determine the level of additives present.

Testing should be done with several doses including one level likely to yield a
mum tumor incidence.

Test numbers of animals currently used are capable only of determining incidences
would not be acceptable if applied to human populations. It should be the objec-
of future testing to move as rapidly as possible to more nearly adequate test num-
recognizing that the number of test animals required to achieve ideal protection is
to exceed the bounds of practicability.

Although it is possible in principle to estimate "safe" levels of a carcinogen, uncer-
es involved in downward extrapolation from test levels will usually result in
sible levels that are the practical equivalent of zero.

The potentiation of carcinogenic activity by dietary cocarcinogens or promoters
be an important public health problem. Incomplete knowledge at this time makes
practical to recommend that routine tests for cocarcinogenicity or promoting
n be conducted. Where circumstances suggest such a possibility, appropriate tests

urged to investigate its copyright status.

should be carried out. Meanwhile, research should be undertaken aimed at determining whether such tests are needed and, if so, how they should be conducted.

11. The use of animals with a substantial tumor incidence of known viral origin (for example, Akr lymphoma) is not recommended for *routine* carcinogenicity studies at this time. A viral profile of the test animals should be obtained.

12. Oral administration should be used for all routine testing of food additives. In special cases, other routes of administration may also be needed.

13. Every effort should be made to assure as "clean" an environment as possible for the animals under study.

14. Ideally carcinogenesis tests should be begun prior to conception and be continued in the offspring.

15. Newborn animals should be reserved for special tests.

16. At the present time there is not enough information available to provide a basis for recommending any rapid test for carcinogenicity.

Implementation of these recommendations will require additional funding from both industry and government for facilities and personnel considerably in excess of that currently available.

REFERENCES

ALLISON, J. B., WANNEMACHER, R. W., HILF, R., MIGLIARESE, J. F., and CROSSLEY, M. L. (1954). Dietary protein and tumor-host relationships in the rat. *J. Nutr.* **54**, 593.

ALLISON, J. B., WASE, A. W., LEATHEM, J. H., and WAINIO, W. (1950). Some effects of 2-acetylaminofluorene on the dog. *Cancer Res.* **10**, 266.

BERWALD, Y., and SACHS, L. (1963). *In vitro* transformation with chemical carcinogens. *Nature (London)* **200**, 1182–1184.

CORNFIELD, J. (1954). Measurement and comparison of toxicities: the quantal response. In *Statistics and Mathematics in Biology*. (O. Kempthorne, T. A. Bancroft, J. W. Gowen, and J. L. Lush, eds.), pp. 327–344. Iowa, State College Press, Ames, Iowa.

CORNFIELD, J., HALPERIN, M., and MOORE, F. (1956). Some statistical aspects of safety tests of Salk poliomyelitis vaccine. *Pub. Health Rep.* **71**, 1045–1056.

DELLA PORTA, G., CAPITANO, J., PARMI, L., and E COLNAGHI, I. (1967). Cancerogenesi da uretano in topi neonati, lattanti e adulti, dei ceppi C57BL, C3H, BC3F1, C3HF1 SAT. *Tumori* **53**, 81–102.

DRUCKREY, H. (1967). Quantitative aspects in chemical carcinogenesis. U.I.C.C. Monograph No. 7, Potential Carcinogenic Hazards From Drugs, pp. 60–77.

DRUCKREY, H., PREUSSMANN, R., IVANKOVIC, S., and SCHMAHL, D. (1967). Organotrope carcinogene Wirkungen bei 65 verschiedenen *N*-Nitroso-Verbindungen an BD-Ratten. *Z. Krebsforsch.* **69**, 103–201.

DURAN-REYNALS, F. (1952). Studies on the combined effect of fowl pox virus and 3-methylcholanthrene in chickens. *Ann. N. Y. Acad. Sci.* **54**, 977–991.

FRIEDMAN, L., MICKELSEN, O., YANG, M. G., WOGAN, G. N., CAMPBELL, J. A., and MORRIS, A. B. (1966). Nutrition Society Symposium, "Nutritional Significance of the Non-Nutrient Components of Food." *Fed. Proc. Fed. Amer. Soc. Exp. Biol.* **25**, 102–144.

GELBOIN, H. V. (1967). Carcinogens, enzyme induction, and gene action. *Advan. Cancer Res.* **10**, 1–81.

HALVER, J. *et al.* (1965). Aflatoxin-induced liver tumours in rainbow trout (Salmo gairdneri). In: *Mycotoxins in Foodstuffs* (Wogan, G. N., ed.), pp. 209–234. MIT Press, Cambridge, Massachusetts.

HEIDELBERGER, C., IYPE, P. T., RÖLLER, M.-R., and CHEN, T. T. (1968). Studies on hydrocarbon carcinogenesis in organ and cell culture. In: *Proliferation and Spread of Neoplastic Cells*. The University of Texas M.D. Anderson Hospital and Tumor Institute, at Houston, pp. 137–156. Williams and Wilkins, Baltimore, Maryland.

...o, L., Laoufur, G. L., and Spatz, M. (1968). Tumor induction in Fischer and ...ne-Mendel rats by a single administration of cycasin. *J. Nat. Cancer Inst.* 41, 1003–...

...ach, S., Druckrey, H., Preussmann, R. (1966). Erzeugung neurogener Tumoren bei Nachkommen nach einmaliger Injektion von Äthylnitroshamstoff an schwangeren ... *Naturwissenschaften B* 53, 410.

...ham, A., Strong, L. C., and Gardner, W. U. (1940). Influence of methylcholanthrene ...cidence of leukemia in several strains of mice. *Proc. Soc. Exp. Biol. Med.* 45, 287–...

... P., and Wisely, A. V. (1963). Production of lung cancer in mice by inhalation exposure ...fluenza virus and aerosols of hydrocarbons. *Progr. Exp. Tumor Res.* 3, 186–215.

...ose, A., Buu-Hoi, and Rudali, Y. (1945). Inhibition of the carcinogenic action pro...d by a weakly active carcinogenic hydrocarbon on a highly active carcinogenic hydro...n. *Brit. J. Exp. Pathol.* 26, 5–12.

...ein, M., and Kaplan, H. S. (1959). Leukemogenic activity of filtrates from radiation-...ed lymphoid tumors of mice. *Science* 130, 387–388.

...y, W., and Saffiotti, U. (1965). Relationships between structure and skin tumorigenic ...y among hydrogenated derivatives of several polycyclic aromatic hydrocarbons. *Ann. Dermatol. Clin. Sper.* 19, 34–44.

...ti, N. (1963). The concept of threshold in carcinogenesis. *Clin. Pharmacol. Therap.* 4, ...100.

...l, N., and Bryan, W. R. (1961). "Safety" testing of carcinogenic agents. *J. Nat. Cancer ...* 27, 455–470.

...ent, P. M. (1965). Carcinogenicity of Aflatoxin-Contaminated Peanut Meals. In: *...toxins in Foodstuffs*, pp. 187–208. MIT Press, Cambridge, Massachusetts.

...ng, H. G. (1966). Feed and foods as sources of carcinogenic factors. *Nutr. Rev.* 24, ...224.

...dgins, H. L., Stier, A. R., and Borsos-Nacht-Nebel, E. (1952). Liver tumor inhibition ...adrenal histologic responses in rats to which 3-methyl-4-dimethylaminoazobenzene and ...ethylcholanthrene were simultaneously administered. *Cancer Res.* 12, 356–361.

...ti, U., and Shubik, P. (1956). The effects of low concentrations of carcinogen in epi...al carcinogenesis. A comparison with promoting agents. *J. Nat. Cancer Inst.* 16, ...969.

...ant, R. (1965). Toxicology of natural products. *Food Cosmet. Toxicol.* 3, 609–620.

...n, M. J. (1938). Studies in carcinogenesis. V. Methyl derivatives of 1,2-benzanthracene. ...r J. Cancer* 33, 499–537.

...s, M. B., Weisburger, J. M., Weisburger, E. K., Gubareff, N., and Suntzeff, V. ... Bioassay of 29 alkylating chemicals by the pulmonary tumor response in strain A mice. ...l Cancer Inst.* 26, 915–935.

...enbaum, A., and Silverstone, H. (1953). Nutrition in relation to cancer. *Advan. Cancer ...* 1, 451.

...ghi, B., Shubik, P., and Della Porta, G. (1960). A Study of skin carcinogenesis in the ...se with single applications of 9,10-dimethyl-1,2-benzanthracene at different dosages. ...er Res.* 20, 1538–1541.

...ghi, B., Magee, P. N., and Barnes, J. M. (1967). Hepatic pathology in rats on low ...ary levels of dimethylnitrosamine. *Brit. J. Cancer* 21, 559–565.

... B. (1963). Development of malignant lymphomas by cell-free filtrates prepared from a ...ically induced mouse lymphoma. *Proc. Soc. Exp. Biol. Med.* 112, 873–875. (1963).

... B., and Shubik, P. (1963). Carcinogenesis in Lewis rats injected at birth with 7,12-...ethylbenz[a]anthracene. *Brit. J. Cancer* 17, 540–545.

... B., and Shubik, P. (1967). Carcinogenesis in AKR mice injected at birth with benzo-...pyrene and dimethylnitrosamine. *Cancer Res.* 27, 43–51.

... B., Magee, P. N., and Shubik, P. (1964). Carcinogenesis study with dimethylnitros-...ne administered orally to adult and subcutaneously to newborn BALB/c mice. *Cancer ...* 24, 1712–1721.

WOGAN, G. N., and NEWBERNE, P. M. (1967). Dose-response characteristics of aflatoxin carcinogenesis in the rat. *Cancer Res.* 27, 2370–2376.

WOGAN, G. N., and NEWBERNE, P. M. (1968). Sequential morphologic changes in aflatoxin carcinogenesis in the rat. *Cancer Res.* 28, 770–781.

VESSELINOVITCH, S. D., and MIHAILOVICH, N. (1968). The development of neurogenic neoplasms, embryonal kidney tumors, Harderian gland adenomas, Anitschkow cell sarcomas of the heart, and other neoplasms in urethan-treated newborn rats. *Cancer Res.* 28, 888 ...

...OLOGY AND APPLIED PHA...

SE

Toxicity and Excr...
Ochratoxin

Toxicity and I...
Ochratoxin A or ...
MOODIE, C. A., ...
Toxicol. Appl. Ph...
pure ochratoxin ...
present in barley ...
anorexia and had ...
of the intubated ...
urine and feces, ...
or kidneys.

...hratoxin A is a myc...
...n der Merwe *et al.* ...
...9). Reports (Krogh ...
...ons in animals fed g...
these feeds is due to ...
nephrotoxins citrini...
...rvations led us to ...
...h pure ochratoxin A ...
...hratoxin A-producin...
...d kidneys and its exc...
Barley cultures B-1 ...
...n Walbeek *et al.*, 1...
...-moldy barley (NM...
...ratoxin A and 1-2 ...
al., 1970) and oxalic ...
...ochratoxin A. For ...
...ubated daily with ...
...rbonate and fed NM...
...MB ad libitum or in ...
...NMB + T groups ...
...ct ochratoxin A at ...
...8) in 24-hr samples ...
...ter (55:45, v/v) extra...
...thanol (4:1, v/v) ext...
...mpounds. Routine h...
...mals necropsied on ...
The average total a...