

# Issues in Biochemical Applications to Risk Assessment: How Should the MTD Be Selected for Chronic Bioassays?

by R. J. Kociba\*

## Introduction

The topic assigned to me was phrased in the form of a question that read, "How should the MTD be selected for chronic bioassays?" As a prelude, the response to the question warrants a brief review of the historical evolution of this issue.

Table 1 is a summary of the historical chronology by which the dose selection process has evolved over the past 30 years. In the early to mid-1950s, the major emphasis was in the evaluation of chronic organ toxicity rather than carcinogenicity, and estimates of potential human exposure were factored into the dose selection process for the animal studies. In the early to mid-1960s, the emphasis shifted to the use of exaggerated (sometimes lethal) doses for the purpose of a short-term qualitative screening evaluation of carcinogenicity potential. This time period was characterized by the use of a relatively crude basis for dose selection, with the emphasis on mortality and frank body weight loss. In the early to mid-1970s, the availability of laboratory animals with better survival led to the lengthening of the duration of dosing with the maximal dose levels compatible with survival of a sufficient number of animals. This time period saw some slight improvements in the relatively crude basis for dose selection.

In the early to mid-1980s, there was a carryover of some of the previous philosophy, but with a recognition of the need for a more scientifically valid basis for dose selection. This time period has been characterized by the need for (a) more definitive subchronic studies, (b) definition and recognition of the critical role played by dose-related changes in kinetics, and (c) the recognition of the need to avoid dose levels that create nonphysiological conditions of treatment. Thus, it is apparent that this issue of dose selection for chronic toxicity and oncogenicity studies has been and continues to be a topic of debate and controversy. I believe this issue of dose selections remains the most challenging aspect of study design, and its importance is underscored by the current

Table 1. Summary chronology of high-dose selection process.

Time period	Guideline for high-dose selection
Early to mid 1950s	Major emphasis on chronic toxicity rather than carcinogenicity Estimates of human exposure factored into procedures
Early to mid 1960s	Use of exaggerated (lethal) doses for qualitative screening of carcinogenicity potential
Early to mid 1970s	Crude basis for dose selection Lengthening of duration of dosing with maximal doses compatible with long-term survival of sufficient numbers of animals Slight improvements in crude basis for dose selection
Early to mid 1980s	Carryover of previous philosophy, but recognition of need for more scientifically valid basis of dose selection More definitive subchronic studies Critical role of dose-related kinetics Avoidance of nonphysiological conditions of treatment Recognition of the need for relevance to known or anticipated levels of human exposure

format of addressing both chronic toxicity and oncogenicity potential in one and the same joint study. Based on the consideration of the extensive human and physical resources that must be appropriated for the conduct of these long-term toxicity studies, it is imperative that we strive to maximize the yield from these efforts by conducting the most scientifically sound type of studies specifically designed to jointly address both chronic toxicity and oncogenicity. The subsequent interpretation of the study results is also greatly facilitated by the use of an optimal joint study design wherein it is more feasible to evaluate any possible mechanistic relationships between chronic toxicity and observed oncogenic responses. Based on these factors, it is recommended that future study designs continue to jointly address both chronic toxicity and oncogenic potential.

There is considerable merit in restating the basic premise and rationale that serves as the driving force to warrant the conduct of these studies. Simply stated,

\*Mammalian and Environmental Toxicology, The Dow Chemical Company, Midland, MI 48674.

these long-term animal studies are conducted to generate both qualitative and quantitative dose-response data that are useful via extrapolation to the evaluation of potential human risk associated with known or anticipated level(s) of exposure by humans. As stated, this definition of the basic rationale stresses not only the qualitative and quantitative aspects of the animal dose-response data, but also the known or anticipated levels of exposure by humans.

Figure 1 graphically depicts the spectrum of biologic responses typically defined in a subchronic or chronic animal toxicity study. The spectrum of biologic response typically spans a graded series of increasing dose levels that ranges from a lower dose level at which no response is elicited (the No-Observed-Effect-Level, or NOEL), to a slightly higher dose level at which we may observe an adaptive physiologic response (the No-Observed-Adverse-Effect-Level, or NOAEL), to increasingly higher dose levels that will define the Lowest-Observed-Effect-Level (LOEL) and also the Frank-Effect-Level (FEL). As depicted in the lower portion of Figure 1, the high dose selected for the chronic toxicity study typically represents a point selected along the range of those dose levels extending from LOEL-LOAEL-FEL. The specific points along this spectrum that will be selected as the high-dose level remains a point of controversy and debate among toxicologists and pathologists. Some scientists recommend selection of a maximally tolerated dose or MTD that is in the range of the FEL, based on the premise that the principal limiting factor should be based on survival of an adequate number of test animals to allow statistical analyses of the data. Other scientists recommend selection of a high-dose level that is in the range of the LOEL or LOAEL, based on the premise that this type of dose selection more adequately generates data that is more readily applicable to the low-dose extrapolation process for human exposure.

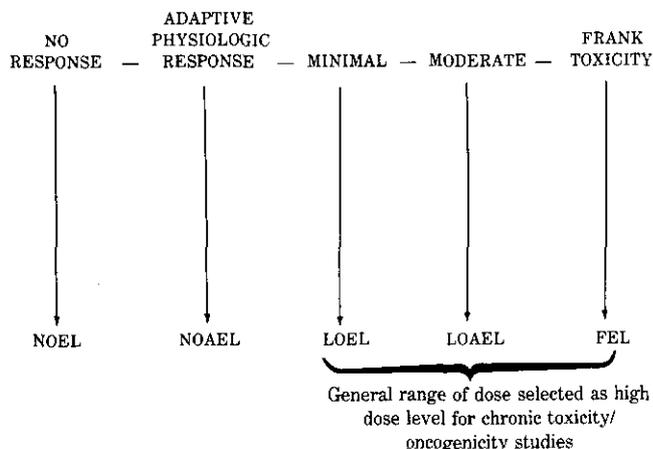


FIGURE 1. Spectrum of biologic responses elicited in subchronic/chronic toxicity studies. NOEL, No-Observed-Effect Level; NOAEL, No-Observed-Adverse-Effect Level; LOEL, Lowest-Observed-Effect Level; LOAEL, Lowest-Observed-Adverse-Effect Level; FEL, Frank Effect Level.

Table 2 lists the perceived advantages and disadvantages currently associated with the selection of maximal tolerated doses for chronic animal bioassays. Proponents of high-dose testing use the primary argument that these bioassays are relatively insensitive for detecting carcinogenic effects. Other scientists express concern over the disadvantages of MTD testing, such as metabolic overload and lack of relevance for safety assessment of human exposure.

Table 3 depicts the key factors that are recommended as the basis for my response to the question of high-dose selection for chronic animal bioassays. Conceptually, it is imperative that the discipline of toxicology strive to utilize the best available technology of the 1980s for high-dose selection. Operationally, this will require that the parameters evaluated in the subchronic (and chronic) dose-response studies must be sufficiently comprehensive to supplement the relatively crude parameters used historically in high-dose selection. Based on conventional clinical and morphologic parameters, the high dose level selected for the chronic studies

Table 2. Examples of current debate concerning the use of maximum tolerated doses (MTD) for chronic animal bioassays.

Perceived advantages of MTD testing	Perceived disadvantages of MTD testing
"The primary argument for support of MTD testing is that current testing protocols (including the MTD) are known to be relatively insensitive for detecting carcinogenic effects" (1).	"Largest doses (far exceeding human exposure levels) has been strongly recommended . . . in order to overcome the inherent low sensitivity of bioassays. These recommendations have been controversial because high doses may themselves produce altered physiologic conditions which can qualitatively affect the induction of malignant tumors" (2).
"More than two-thirds of the carcinogenic effects detected in feeding studies would have been missed had the high dose been reduced from the estimated MTD to 1/2 MTD" (1).	"Further bias is introduced by prescribing 'nearly toxic' or 'maximally tolerated' doses of the test substance, i.e., doses that potentially overload these metabolic pathways that might influence the development of a tumor" (3).
	"To slavishly endorse the MTD is to ignore the only purpose of the study: develop some understanding of risk/benefit at real exposure limits" (4).
	"The use of the MTD is an anachronism" (4).
	"The MTD might be an excessive multitude of the human exposure. We have examples where the MTD is 30,000 × the human exposure" (4).
	"Nonsensical academic exercise . . . Tumors (or anything else) observed at these doses are essentially irrelevant to safety considerations for man" (5).

**Table 3. Recommendations for selection of high dose levels for chronic animal bioassays.**

Conceptually, the discipline of toxicology must strive to utilize the best available technology of the 1980s for high-dose selection

The parameters evaluated in the subchronic (and chronic) dose-response studies must be sufficiently comprehensive to supplement the crude parameters used historically to define the MTD

Detailed histopathology  
Hematology, urinalysis, clinical chemistry  
Organ weights  
Specific tests, dictated by mechanism of action of test compound  
Methemoglobin  
Acetylcholinesterase  
Others

Optimal choice of high dose level for chronic study should be that dose which in subchronic study elicited discernible but slight degree of toxicity

Sufficient metabolism and kinetic studies must be prospectively conducted to allow selection of high dose levels that:

Are within the range of linear dose-dependent kinetics  
Avoid metabolic overdose

Factor in the known or anticipated levels of human exposure

should represent that dose level which elicited some discernible but slight degree of toxicity in the subchronic studies. Good science (and humane reasons) dictates against the selection of higher dose levels for the chronic studies that exceed that dose level which in the subchronic studies has elicited a discernible but slight degree of toxicity.

We must also prospectively conduct sufficient metabolism and kinetic studies that allow selection of high-dose levels that are (a) within the range of linear dose-response kinetics, and (b) avoid metabolic overdosing. Finally, we must factor in the data on the known or anticipated levels of human exposure associated with the substance. It is only through the use of these recommendations that the discipline of toxicology can profess to be applying the best available technology of the 1980s to the issue at hand, namely, selection of high-dose levels that are scientifically appropriate and useful for the subsequent extrapolation to human exposure scenarios.

## Discussion

DR. ROY ALBERT, UNIVERSITY OF CINCINNATI MEDICAL CENTER: I think it's generally recognized that a bioassay is trying to do two things at once, which are in a sense mutually incompatible: first, to answer the question as to whether the agent is a carcinogen, second, to get a handle on the nature of the dose response as a basis for extrapolation to low levels. I wonder whether or not we haven't been drawn into an exceedingly expensive and a nonoptimally productive approach to this by the conventional elaborate bioassay technique. And I wonder whether or not we ought to roll back the clock a little bit, as you indicated in one of your earlier slides, and do screening studies for carcinogenic response at very high doses for MTD.

Carcinogens are by and large pretty toxic materials.

And if you really want to be sure whether or not you're going to get a carcinogenic response, you generally want to aim for exposures that are really in the toxic range. Having done a relatively abbreviated form of screening for the existence of carcinogenic response, this then could serve as a springboard for the much more elaborate dose-response type of bioassay, which is very costly, and which would provide the basis for making judgments as to how the agent is to be dealt with in a regulatory framework.

The question I'm raising has to do with the overall efficiency of our searching out carcinogens in the environment. It's been said *ad nauseam* that there are in the order of 60,000 chemicals that are in circulation. We have examined perhaps 10% of these agents. Wouldn't it be better to fall back on screening approaches for carcinogens and limit the bioassays to those that show up positive in screening rather than trying to do everything at once at a cost of a half-a-million dollars a throw?

DR. KOCIBA: I have an answer I'd like to give to that, but before I respond I'll give other people in the audience a chance to give their opinion on this very timely question that Dr. Albert has raised.

DR. RAJENDRA CHHABRA, NIEHS: I don't have many comments. I'm just asking about some of your definitions. What is adaptive physiological response compared to no response? How one determines what is adaptive physiological response?

DR. KOCIBA: I'll answer your question first and then we'll come back to Roy's because I think Roy has a very pertinent question. In response to that question, in my opinion, the adaptive type of response has to be obviously looked at on a very specific case-by-case basis. I know of no general rule that I could give you or that anyone could give you that would serve 100% of the time. I'm thinking along the lines of enzyme induction in hepatocytes in the absence of accompanying cytotoxicity and this sort of thing as an adaptive response.

DR. CHHABRA: If we find in one group of animals there's a slight induction of drug-metabolizing enzymes, in another group there isn't, is the one group which has got slight induction of drug-metabolizing enzymes an adaptive physiological response?

DR. KOCIBA: It could be.

DR. CHHABRA: So would that dose be selected as a high dose?

DR. KOCIBA: It may or may not be. Like I say, I know of no general rule or generic rule one could use there. I think you would have to look at the full data base, and therein would lie the advantage of having as comprehensive a data package as you have. This would be not only data from your subchronic studies that would allow you to interpret those observations and look at a dose response, but also use what data we have on kinetics to see is there a break in the linearity of the curve that would correlate with some of these things.

DR. CHHABRA: I agree with you. My second question is what are the nonphysiological conditions of treatment?

DR. KOCIBA: You're questioning some of the terminology used in the literature?

DR. CHHABRA: Yes. I'd just like to make it clear to myself.

DR. KOCIBA: Well, I think what people have been trying to do in the literature is to group together all of those components that come into play as you scale up the doses up into a range in which you have exceeded the normal metabolic capacities inherent within that organism. Somewhere along the line those are going to be exceeded. And I think that's the meaning that some of the people I've cited there have had in mind when they use the term nonphysiologic conditions.

DR. CHHABRA: And anything you use, dose lower than that, that's within physiological conditions?

DR. KOCIBA: Well, that's what your dose response would tell you.

DR. CHHABRA: And one last question. How do you determine this slight degree of toxicity? You said that for selection of MTD you would have a slight degree of toxicity. What is that slight degree of toxicity?

DR. KOCIBA: Again, that's going to depend on your case-by-case evaluation of the data set you have. And you're going to be much better off if you have a comprehensive data package with your complementary clinical data, your organ weight data, your clinical observations, etc., to supplement the morphology. Too many times I've seen the dose selection based entirely on morphology and maybe a body weight reduction. I think those parameters are too crude. In the later 1980s that we are now in I think we have the responsibility to society at large to do a better job than we've done in the past.

DR. CHHABRA: Okay. My point for asking those three questions was I think it would be a subjective decision and it would be a sort of professional judgment on selection of the high dose. There are no well-defined criteria that can help us to determine what is slightly toxic or what a nonphysiologic condition is, and what is adaptation. So it would be mostly a professional judgment in the selection of high dose MTD for carcinogenicity studies, and it will vary from person to person. And from chemical to chemical.

DR. KOCIBA: In my experience there have never been two compounds that were in essence two peas in the same pod. You know, there might be some similarity. But you have to consider each compound on its own merits. And that's why I'm making the plea to have a more comprehensive data package before we go about selecting those dose levels. Typically, the kinetic studies are going to be conducted sooner or later. And my plea is let's do these studies before we make that dose selection rather than 3 years later when we're trying to do it in retrospect.

DR. CHHABRA: But you cannot do kinetic studies on all kinds of chemicals. It depends on the purity of the chemical; a mixture of chemicals cannot be subjected to pharmacokinetic studies.

DR. KOCIBA: Well, if there's that much interest in subjecting it to a study that Dr. Albert says is going to

cost over a half-a-million dollars, I think we have to rethink our priorities in regard to how we allocate available resources for these different studies. Maybe it will amount to testing fewer materials but doing a better job on the ones we do study.

Before we go on to the next question, I'd like to go back to Dr. Albert's very pertinent question here.

DR. MARSHALL ANDERSON, NIEHS: That's what I wanted to go to. As Dr. Albert pointed out, the two purposes are to test whether there's potential of a chemical to be a carcinogen at whatever dose, and extrapolation to low doses. Let's forget about the species extrapolation. I don't care how you set up the dose response; if all you do is count tumors, you'll never do low-dose extrapolation by just counting tumors on the animal. I mean, I think NCTR's EDLI study proves that point. How many chemicals could you test 50,000 animals with? So the only way that I think you're going to scientifically extrapolate to low doses is by incorporating the mechanisms that we know about—at least some of the steps we have pretty well defined. You have to take these into account.

Looking at promutagenic lesions, you can go down to doses to which humans are exposed. I think that should be the approach rather than trying to design tumor studies to get down to human exposure doses. I don't think it's possible.

DR. KOCIBA: Do you have a comment in regard to what Dr. Albert spoke to?

DR. J. C. BHANDARI, DYNAMAC CORPORATION: Yes, please. In response to Dr. Albert's question, a couple of years ago the Society of Toxicologic Pathologists had a symposium which was the design of carcinogenicity studies. And in that one I had proposed the same idea that you just suggested by presenting certain examples. You could actually boost the dose up and save millions and time. But there were a couple of problems with it. An example came very easy. Vinyl chloride. At 1000 mg you can produce all the tumors you want to see at 4 months or you can wait 2 years or 18 months, say, at 50 [mg] or less. Many other drugs can produce good adenomas in livers and so on within 3 months or 6 months depending on your doses and see a beautiful nice correlation. And these are not counting dead also. These are microscopic and convincing results. But there are two problems. One is just because you are at a way high dose there. In the regulatory agencies now their approach is the benefit of doubt. When you have a positive now, you have 1000 or 3000 times of what would be the human dose, as Dr. Kociba showed. Just because you wanted to solve (whether or not it was carcinogenic), you went that 1000 times or 3000 times and now you've got a positive, whereas on the low doses you could show perfect 50 on it. You've got a real problem as far as the industrial point of view. It becomes very difficult to convince the regulatory agencies of that point now that you've got a 1000-fold of 50. So that's one problem right there.

The second problem was how about the negative ones? Positive ones, sure, you could save time and all

that. But the negative ones, I wasn't concerned with it because you got to do a study anyway. So if you can show negative at 1000-fold, fine, you've done your job. So negative really was not a problem. Because negative ones you won't save the time because you still have to go 2 years. You cannot say, well, we'll be done in 6 months or something. But at least you have proven your point.

But I think the biggest problem was the regulatory agencies have to accept that fact. Okay. Fine. If you will show so much for safety, no problem. But that doesn't happen in practice.

DR. KOCIBA: Well, I'll give my opinion in regard to the question that Dr. Albert raised. I think if one would go to a testing scheme where you would segregate out carcinogenicity qualitative assessment from dose response quantitative definition of organ toxicity and so forth, I think you would create a dilemma that I think was alluded to in the previous comments. If you did your qualitative high dose abbreviated scaled-back kind of carcinogenic assessment and that study elicited a positive response, this would automatically be put into the unit risk estimate, and it would almost be a moot point to spend the hundreds of thousands of dollars to do anything in regard to the lower dose testing.

I personally endorse the concept of incorporating the oncogenicity and chronic toxicity into the same test. And the reason for this is multiple. One of which it's only when we have the data from the same sort of study that we can make any association between precursor lesions, cytotoxicity, hyperplastic lesions with the neoplastic process where it was observed. If we were to segregate out these end points, we would lose that distinct option of trying to make our interpretation on the best basis, taking into account what precursor lesions were occurring either at the same dose level that would cause an oncogenic response, or, more importantly, at lower dose levels is to see how steep that dose response is in regard to the biological effect elicited. So personally, I endorse the collaboration of looking at both the chronic toxicity and oncogenicity in one and the same study. And I think when you're all said and done, as long as we continue to do it in this manner using animals and so forth, the cost would probably not be that much different.

DR. BYRON BUTTERWORTH, CHEMICAL INDUSTRY INSTITUTE OF TOXICOLOGY: I think when we're dealing with compounds in the hundreds of milligrams per kilogram per day region that the issues you've raised are reasonable, and we all have to stick with what we're dealing with. I'm particularly concerned, however, with that very small subset of chemicals where people go to massive doses in order to get the tumors. I think we're wasting resources and we're losing credibility. For example, in the pharmaceutical industry, I've been told over and over again that if you're dealing with a nontoxic chemical that's in development you're in real trouble. Because if reaching that maximum tolerated dose required thousands of milligrams per kilogram per day, you invariably encounter some strange problem that's

a result of these unnatural conditions. For example, in the NTP program that Ray [Tennant] talked about and looking at that data that he presented at the Williamsburg meeting, there were several chemicals that were tested at doses in excess of 6000 mg/kg/day. You really have to wonder if we haven't gone off the deep end for those.

I would suggest that people that are involved with this problem (Joe Haseman mentioned that he had a problem with things that are going to be missed at higher and lower doses) if maybe it wouldn't be a good exercise to look at those chemicals which induce tumors only at doses in excess of 1000 mg/kg/day. And then ask the question, If we had chosen a thousand as the cutoff and that would have been judged a carcinogen and we would have done risk assessment based on other toxicological parameters, would human health really have been affected?

We really need to begin to focus on this small subset because it's hurting the rest of the field. The compounds that we deal with that are in the reasonable range, I think we're doing all right. But this subset is really killing us. And once again, I think that many of these are going to be of the nongenotoxic variety. For example, if you look at the NTP correlation study that Ray talked about, those compounds that required in excess of 500 mg/kg/day not a single one was a *Salmonella* mutagen.

Tying this back to Ray's final point, are there second-class citizens? I don't think that he can say that nongenotoxic carcinogens *per se* are less dangerous. Because there's TCDD and many others. However, if you put two parameters together, if you say it's a nongenotoxic carcinogen and it required greater than 1000 mg/kg/day to produce the tumor, then maybe you have a candidate for some second-class citizens. That is to say, chemicals which require less of our attention than maybe some of the others.

DR. MICHAEL DIETER, NIEHS: As one of the chemical managers involved in this program in selecting these doses and trying to carry out these tests, I'd like to make the remark that the other side is also valid. Just as in the genotoxic tests, one of the things that should be avoided is a possibility of a false negative. Just as the example you cited, the chemical was used in very high doses. There's another example that could be cited just as well on the other side of the coin where the chemical was tested in the tenths of parts per million range and found to be completely negative in a 2-year test. And when tested at tens or hundreds of milligrams per kilogram in another test it was found to be carcinogenic, causing multiple tumors in both sexes and both species.

If you look at a scenario of this sort by selecting a dose level, the highest of which only caused some sensitivity responses in the treated animals in the prechronic studies, you would have entirely missed a carcinogenic response by not selecting higher doses.

UNKNOWN SPEAKER: Let me respond. Don't get me wrong. If you are dealing with tens of milligrams per

kilogram per day, surely you must deal with the issues the way we have been dealing with them. I agree with you. I'm talking about that set that's way over there where you need over 1000 mg/kg/day. And then I challenge you to show me that we're dealing with the real problem.

So it's not the ones that are in the range of, say, less than a thousand. It's the ones greater than a thousand where I think we're losing credibility and we're wasting resources.

DR. KOCIBA: Dr. Albert has motioned that we must move on to the next paper. I want to thank everybody for participating. I'm sure we could go on for 3 hr on this issue.

## REFERENCES

1. Haseman, J. K. Issues in carcinogenicity testing: Dose selection. *Fundam. Appl Toxicol.* 5: 66-78 (1985).
2. U.S. Office of Science and Technology Policy. Chemical carcinogens: A review of the science and its associated principles. *Federal Register* 50: 10372-10442 (1985).
3. American Medical Association Council on Scientific Affairs. Carcinogen regulation. *J. Am. Med. Assoc.* 246: 253-256 (1981).
4. Young, S. S., and Brannon, D. R. Letter to the editor. *Fundam. Appl. Toxicol.* 6: 185-186 (1986).
5. Meyers, D. B., Young, S. S., and Gries, C. L. Letter to the editor. *J. Natl. Cancer Inst.* 74: 1151-1152 (1985).