12-07-07

Analysis of in vivo Mutation Data Can Inform Cancer Risk Assessment

Martha M. Moore[1*], Robert H. Heflich[1], Lynne T. Haber[2], Bruce C. Allen[3], Annette M. Shipp[4], Ralph L. Kodell[1,5]

[1] National Center for Toxicological Research, Food and Drug Administration, Jefferson, AR

[2] Toxicology Excellence for Risk Assessment, Cincinnati, OH

[3] Bruce Allen Consulting, Chapel Hill, NC

[4] ENVIRON Health Sciences Institute, Ruston, LA

[5] Present address, Department of Biostatistics, University of Arkansas for Medical Sciences, Little Rock, AR

* Corresponding author: Martha M. Moore, National Center for Toxicological Research, HFT-120, 3900 NCTR Road, Jefferson, AR 72079, Tel: 870-543-7050, Fax: 870-543-7393, Email: martha.moore@fda.hhs.gov

Short Title: Mutation data informs cancer MOA

Word count:

Abstract:

Text:

References:

Disclaimer: The views presented in this article do not necessarily reflect those of the U.S. Food and Drug Administration.

Acknowledgement: The authors would like to acknowledge Michael L. Dourson for his significant participation and contributions during the development of the concepts described in this paper.

ABSTRACT

Under the new U.S. Environmental Protection Agency (EPA) Cancer Risk Assessment Guidelines (U.S. EPA, 2005), the quantitative model chosen for cancer risk assessment is based on the mode-of-action (MOA) of the chemical under consideration. In particular, the risk assessment model depends on whether or not the chemical causes tumors through a direct DNA-reactive mechanism. It is assumed that direct DNA-reactive carcinogens initiate carcinogenesis by inducing mutations and have low-dose linear dose-response curves, whereas carcinogens that operate through a nonmutagenic MOA may have nonlinear dose-responses. We are currently evaluating whether the analysis of in vivo gene mutation data can inform the risk assessment process by better defining the MOA for cancer and thus influencing the choice of the low-dose extrapolation model. This assessment includes both a temporal analysis of mutation induction and a dose-response concordance analysis of mutation with tumor incidence. Our analysis of published data on riddelliine in rats and dichloroacetic acid in mice indicates that our approach has merit. We propose an experimental design and graphical analysis that allows for assessing time-to-mutation and dose-response concordance, thereby optimizing the potential for in vivo mutation data to inform the choice of the quantitative model used in cancer risk assessment.

Key Words: Cancer risk assessment, mode-of-action, mutagenic carcinogen, nonmutagenic carcinogen

INTRODUCTION

Cancer risk assessment uses mode-of-action (MOA), defined as the sequence of key events leading to cancer (Cohen *et al.*, 2003; U.S. EPA, 2005), for both qualitative and quantitative data analysis. Cancer often involves two basic key events, mutation and cell proliferation (Hanahan and Weinberg, 2000), the timing of which is critical to defining the MOA for a chemical carcinogen (e.g., Meek *et al.*, 2003). Chemicals can induce cancer by inducing mutations as a direct result of their reaction with DNA, but a variety of other MOAs are possible. For instance, chemicals can cause increased cell proliferation by acting directly as a mitogen, or by mechanisms such as hormone disruption, epigenetic alterations of cell cycle control mechanisms, or by binding to cellular receptors and altering signal transduction. Increased cell proliferation may result in tumors by facilitating the expansion of preexisting mutations, or may result in tumors by increasing the frequency of "spontaneous" mutations, for example by decreasing the time for repair of endogenously damaged DNA (Melnick *et al.*, 1993; Thilly, 2003).

Carcinogenic chemicals that are DNA-reactive and, as a direct result of this interaction, induce mutations that are causally involved in cancer, are defined as "mutagenic" carcinogens. Nonmutagenic carcinogens also may cause mutations, but these mutations are produced as a consequence of other key events induced by the agent and are not a direct result of the interaction of the agent with DNA. For regulatory decision-making, distinguishing between chemicals that cause cancer via either a mutagenic or a nonmutagenic MOA is a key decision point for the dose-response evaluation. In particular, under the U.S. Environmental Protection Agency's (EPA's) 2005 Cancer Guidelines, a linear approach is used for low-dose extrapolation

for mutagenic carcinogens, while a nonlinear approach may be used for nonmutagenic carcinogens (U.S. EPA, 2005).

Much of the data used for establishing MOA (e.g., measurements of apoptosis and cell proliferation) is generated in the rodent tumor target tissue(s) (Meek *et al.*, 2003). In contrast, most determinations as to whether a chemical has a mutagenic cancer MOA are currently based on a weight-of-evidence (WOE) assessment of data from a number of test systems, many of which are in the vitro assays used for hazard identification (Dearfield and Moore, 2005; Cimino, 2006). Although these assays provide useful information, it is necessary to extrapolate from these assays to address the key questions regarding MOA. To determine whether or not a chemical acts via a mutagenic MOA, one would ideally determine not only whether the chemical is a mutagen, but whether it causes gene mutations in the target tissue for tumorigenesis, and whether those mutations play a causal role in cancer development. To address these latter two considerations, we propose conducting experiments designed to evaluate the dose-response and temporal development of in vivo mutations in the tumor target. The design also incorporates the evaluation of premutational events (DNA adducts and/or oxidative damage) and other biological effects (inflammation, cell proliferation and/or cytotoxicity), as well as an evaluation of the temporal association of these effects with the induction of mutations (if any) and the induction of preneoplastic lesions if such lesions occur prior to the formation of frank tumors. All of the various endpoints should be evaluated in such a way as to obtain dose-response information that can be used to create a time-line for the occurrence of the events and the dose-response concordance of the events.

This holistic evaluation of the various effects following chemical exposure and the possible induction of mutation in the target tissue optimizes the information available for the MOA assessment. We feel that this approach provides a more scientifically rigorous assessment as to whether chemicals that are positive in the standard genetic toxicology battery and other classical mutation assays are capable of inducing mutation by direct DNA reactivity in the tumor target tissue, and whether these mutations play a causal role in tumorigenesis. The approach also expands upon current approaches which evaluate primary DNA damage (DNA adducts and/or oxidative damage) but do not determine whether, in fact, mutations are actually induced in the target tissue.

Before going further, we need to stress that the most scientifically rigorous approach for determining MOA is to assess all of the possible biological effects caused by the chemical treatment and that the most appropriate genetic targets to address the role of induced mutation in such assessments would be the oncogenes and the tumor suppressor genes involved in each specific tumor type. It also is desirable to use these data to develop complete biologically-based dose-response (BBDR) models for conducting quantitative risk assessments for individual chemicals. However, in many (most) cases the resources required for such extensive experimental efforts are not available and therefore regulatory decisions are often made in the absence of a BBDR model. We feel that our approach is a reasonable compromise between obtaining a full BBDR and the current practice of conducting a weight-of-the evidence evaluation using data collected from a number of different experimental designs and without the benefit of knowing whether, in fact, the chemical in question is actually capable of inducing mutation in the target tissue.

Our manuscript is intended to outline a new strategy for using in vivo mutation data to inform cancer risk assessment. To that end, we have developed statistical approaches to use in vivo mutation data in the tumor target tissue as a part of the WOE for MOA assessment. We have written this manuscript to stimulate thinking and discussion in the regulatory community. Because our approach is new, the exact experiments that we outline have not been undertaken and, therefore, the optimal data for assessing our approach are not available. We will, however, utilize some existing data to demonstrate the general concepts of our approach.

As already indicated, our strategy involves assessing the ability of a chemical to induce gene mutations (base pair substitutions, frameshifts, deletions and insertions that affect the function of a specific gene) in the tumor target tissue under the same treatment conditions used to induce tumors. Transgenic shuttle vector models currently provide the best approach for measuring in vivo gene mutations. In these models, gene mutations are detected using transgenes, such as *lacI*, *cII*, or *gpt* delta, that are incorporated into the nuclear DNA of every cell of the animal (Thybaud *et al.*, 2003). Unlike oncogene/tumor suppressor gene mutations, the shuttle vector mutations do not confer a differential growth phenotype to the cells (i.e., the mutations are genetically neutral) (Heddle *et al.*, 2003), and these mutations can accumulate with time following chronic exposures (Zhang *et al.*, 1996; Heddle *et al.*, 2000; Chen *et al.*, 2001), allowing a cumulative measure of mutant frequency (MF). Although the targets for mutation in these models are bacterial genes and mutation detection is performed by rescue of the transgene and growth in indicator bacteria, extensive studies indicate that the mutations detected with these systems are induced in the animal, rather than being generated in the indicator bacteria (Hill *et*

al., 1999). Also, validation studies indicate that the frequency and spectrum of mutations detected in the transgenes are a good approximation of mutations in endogenous genes, particularly for base pair substitutions and relatively small deletions and insertions (Skopek *et al.*, 1995; Okongi *et al.*, 2001; Chen *et al.*, 2001).

Chromosomal mutations (multi-gene deletions, translocations, mitotic recombination, and aneuploidy; events involving more than a single gene and often visible cytogenetically) are important in the etiology of tumors, particularly because they can induce loss of heterozygosity, a major factor in tumor development (Knudson, 2002). Hazard identification strategies evaluating the ability of chemicals to cause genetic damage include the detection of both gene mutations and chromosomal mutations (Dearfield *et al.*, 1991; Cimino, 2006). However, in the selection of a quantitative model for cancer risk assessment, it is important to distinguish between gene mutation and chromosomal mutation as a key event. The induction of gene mutations is thought to occur with linear kinetics (one-hit model), while the induction of chromosomal mutations often occurs in two steps (e.g., translocations are produced by chromosome breakage followed by joining to produce a new arrangement of chromosomal material) and are thought to occur with nonlinear kinetics [two- (or more) hit model]. The most commonly used transgenes (e.g.. *lacZ*, *lacI,* and *gpt*) can be used to detect gene mutations, but not chromosomal mutations (Thybaud *et al.*, 2003). Thus, while the inability of transgenic mutation assays to detect chromosomal mutations is a disadvantage for hazard identification, it is a benefit for informing the MOA determination (and implications for low-dose extrapolation) for cancer risk assessment.

Under the U.S. EPA's 2005 Cancer Risk Assessment Guidelines (U.S. EPA, 2005), determination of whether a chemical acts via a mutagenic MOA is important both for informing the approach to low-dose extrapolation (where the default is to assume that the chemical acts via a mutagenic MOA), and for applying Age-Dependent Adjustment Factors (ADAFs) for early life risk (where the default is to assume that the chemical does not act via a mutagenic MOA). This determination follows the modified Hill criteria, and evaluates issues such as the strength, consistency, and specificity of association, dose-response relationship, temporal relationship, and biological plausibility and coherence. To expand upon its guidance, EPA has recently released in the Federal Register for public comment a "Framework for Determining a Mutagenic Mode of Action for Carcinogenicity" (http://www.epa.gov/osa/mmoaframework/pdfs/MMOA-ERD-FINAL-83007.pdf.). The study design presented in this article specifically addresses whether data for gene mutations in the target tissue are consistent with mutations being a key event for tumorigenicity, based on the criteria of temporality and dose-response concordance as outlined in EPA's framework.

Many of the high-profile chemicals currently being evaluated for human cancer risk (e.g., acrylamide, trichloroethylene, acrylonitrile) are either mutagens or are metabolized to chemicals that can induce mutation. It is not clear, however, whether or not these chemicals are mutagenic carcinogens; that is, it is not clear whether mutation is the key event for their tumorigenicity in animal models. Indeed, there have been numerous suggestions in the literature that nonmutagenic MOAs are responsible for, or play a primary role in, the carcinogenicity of these chemicals (e.g., Shipp et al., 2006; Clewell and Andersen, 2004; Moore and Harrington-Brock, 2000; Kirman et al., 2005). However, using current methods for weighing the available genetic

toxicology data, even weak evidence for mutation via direct DNA reactivity would almost always result in the presumption of a mutagenic MOA. (The possibility of multiple MOAs is considered in the Discussion section.) To facilitate the development of a more rigorous approach for determining whether or not a chemical acts via a mutagenic MOA, we have evaluated the feasibility of using in vivo gene mutation data obtained from tumor target tissues. Specifically, we propose using in vivo mutation data and the modified Hill criteria of temporality and dose-response concordance in evaluating the MOA for tumor induction. It is our intent to stimulate discussion concerning the development of an optimal experimental design and modeling methods for comparing mutation and cancer data that will best inform the MOA assessment.

PREDICTIONS

Temporality

There are multiple steps (stages) in cancer development. When mutation is the key initiating event, as would be anticipated for a directly DNA-reactive carcinogen, mutations are induced early in the process, followed by a cascade of additional events resulting in the manifestation of a tumor. There should be an increase in the frequency of gene mutations following relatively short treatment times and prior to the observation of preneoplastic lesions and tumors. On the other hand, if a chemical causes an increase in the number of gene mutations only after an extended period of chronic exposure (likely preceded by toxicity and/or cell proliferation and possibly including clonal expansion, and alterations in gene expression), or no increase in gene mutations even following very long treatment times, this supports a nonmutagenic MOA.

Dose-Response Concordance for Mutation and Tumor Induction

When gene mutation is the key event, the dose-response for mutation should *lead* that for tumors. That is, gene mutations will be induced at doses equal to or lower than those required to form tumors. Alternatively, if the dose-response assessment shows a positive response for tumors at doses lower than those required to cause gene mutations, this observation is consistent with a nonmutagenic MOA.

PROPOSED EXPERIMENTAL DESIGN

To our knowledge, no studies have been conducted specifically for evaluating the temporality and dose-response concordance of mutagenicity and carcinogenicity. We propose such an experimental design. To inform temporality, we define a new assessment, time-to-mutation, that capitalizes on the expectation that (1) chemicals that induce gene mutations in a relatively short time following the first exposure are more likely to be mutagenic carcinogens and (2) the finding of no increase in MF or an increase only after relatively prolonged treatment supports the determination that a chemical is a nonmutagenic carcinogen.

Unfortunately, we do not have adequate experimental information to define the time-to-mutation for chemicals acting by a mutagenic MOA versus chemicals acting via a nonmutagenic MOA. Based on very limited data, the recommended design for hazard identification studies using transgenic mutation assays is 28 days of treatment followed by a 3-day expression period (Thybaud *et al.*, 2003). To support our contention that chemicals that are promoters (or other stimulators of cell division) may increase the mutant frequency following long, but not short exposure times, an analysis of several apparently nonmutagenic carcinogens found that oxazepam, phenobarbital, and Wyeth 14,643 induced weak mutagenic responses in Big Blue

mice after 6 months of treatment (Singh et al., 2001). We propose using several mutation sampling points to evaluate time-to-mutation during chronic exposure to the test agent. A reasonable first approach may be to use 1-, 2-, 4-, and 6-month exposures, with the possibility of stopping the exposure after an increase in MF is observed. This timing will be reconsidered as experiments are conducted to better define the optimal exposure times. It should be noted that a negative response for mutant induction in the target tissue, even after extended exposure, provides strong evidence that the tumors are being induced via a nonmutagenic MOA.

For the dose-response concordance evaluation, doses will be selected based on those used in the tumor bioassay. Using 6 or more dose groups, with perhaps fewer than the typical 5-6 animals per group, will enhance the precision with which the overall dose-response curve can be estimated (Kavlock *et al.*, 1996). Doses below those that were positive in the cancer bioassay will be included for the evaluation to increase the accuracy of dose-response assessment in the low-dose region. To facilitate direct comparison of the data, the exposure conditions will be designed to match the conditions under which the positive cancer data were obtained. Ideally, the experiments will be performed in the same species, using the same exposure conditions, and assaying the same tissues (and perhaps types of cells) in which tumors were observed. Both the Big Blue and delta *gpt* assays can be conducted in F344 rats and B6C3F1 mice (Dycaico *et al.*, 1994; Thybaud *et al.*, 2003; Hayashi *et al.*, 2003), making it likely that transgenic rodents can be used with the same genetic backgrounds employed for the tumor study.

Experiments will be designed individually for each test substance and will be based on the information available for that chemical. Often, information concerning the types of DNA

adducts that are produced in the target or other tissues in vivo will be available already. There may be information concerning the ability of the chemical to induce oxidative damage or cytotoxicity, cell proliferation and inflammation. Some chemicals (such as DCA) have already been extensively evaluated for preneoplastic lesions and/or other pathologically observable effects. The final design for a specific chemical will incorporate all of the relevant possible endpoints and will be conducted so each of the endpoints can not only be placed on a timeline incorporating mutation, but also so that a dose-response evaluation can provide insight into what effects occur at what exposure time and at what dose. Because the time to mutation or other effects will certainly be influenced by the dose, it is important that there are sufficient data to fully assess temporality and dose response for every endpoint. For a mutagenic MOA to be established, it must be clear that the induction of mutation occurs prior to the induction of tumors (or preneoplastic lesions) at each dose.

Dose-response analysis will employ all of the data (including mutation) obtained at each time point and will be used for assessing temporality. Dose-response analysis combined with time to the induction of primary DNA damage, time-to-mutation, and time to the induction of the other relevant biological effects should provide information as to whether mutation is likely to be *the* key event in the etiology of specific target site tumors. These targeted mechanistic studies, together with toxicokinetic data and other information, could contribute to the overall WOE determination for cancer MOA. Mutational spectrum analysis may provide additional information as to whether any small increase in MF is, in fact, indicative of chemical-induced mutations or an expansion of preexisting mutations.

CASE STUDIES

While studies following the above design have not yet been conducted, the general concept can be illustrated by analyses of published gene mutation and tumor data. Unfortunately, these studies lack many of the key features of the design presented above, and should be thought of more as exploring the concept rather than a full demonstration of the approach. We present two case studies, one (riddelliine) that we believe is consistent with the chemical being a mutagenic carcinogen and one (dichloroacetic acid) that is more consistent with the chemical being a nonmutagenic carcinogen.

Riddelliine

Riddelliine is a naturally occurring, mutagenic pyrrolizidine alkaloid that can contaminate human food (reviewed in Fu *et al.*, 2004). For our analysis, we used liver mutation data from Big Blue rats exposed for 12 weeks (Mei *et al.*, 2004a). MF is expressed as the number of mutant lambda phages per 1 x 10^6 phages recovered from the livers of rats. In the Big Blue F344 rat model each recoverable phage has a single mutational target gene (in this case the *cII* gene), and each cell contains approximately 20 phages (Dycaico *et al.*, 1994), providing multiple cellular targets for mutations. The cancer data were obtained from a National Toxicology Program (NTP) two-year carcinogenicity study (NTP, 2003; Chan *et al.*, 2003). A weakness in the analysis of these data sets is that the tumors from the NTP study are hemangiosarcomas and arise in the liver vascular tissue while the mutation analysis was conducted using whole liver. Our approach will consider examining mutations in particular cell types where feasible and where the tumor data indicate that particular cell types are involved. A study of riddelliine mutagenesis in rat liver indicates that mutation induction is much higher in endothelial than parenchymal cells (Mei *et*

al., 2004b); thus the MF measured in the whole liver may have been an underestimate of the MF in liver vascular cells. The MF and cancer data are summarized in Table 1.

Dichloroacetic Acid (DCA)

DCA is a mutagenic drinking water disinfection by-product (Moore and Harrington-Brock, 2000). The mutation data used for our analysis came from the liver of Big Blue mice exposed to DCA via their drinking water for 4, 10, and 60 weeks (Leavitt *et al.*, 1997) and the cancer data are from DeAngelo *et al.*, (1999).[1] The MF and cancer data are summarized in Table 2.

MATERIALS AND METHODS

Time-to-mutation

We plotted the 3 time points (4-, 10-, and 60-week exposures) for DCA MF (Figure 1a). Unfortunately, there is only one time point for riddelliine MF (a 12-week exposure). To provide some comparison between these two chemicals, we plotted the riddelliine MF at 12 weeks of exposure with the DCA MF at 10 weeks of exposure (Figure 1b). For both comparisons, we plotted the induced MF (the observed frequency of mutant transgenes less the background MF) to better visualize the actual increase in MF above the concurrent background (which varies with age and animal model).

Dose-Response Concordance

Although the EPA Cancer Guidelines identify dose-response concordance as a major part of the MOA analysis, there is currently no recognized optimal approach for comparing dose-responses

[1] We thank Dr. Anthony DeAngelo for providing the primary liver carcinoma data from his bioassay of mice exposed to DCA via their drinking water (published in DeAngelo *et al.*, 1999).

generated by different biological measurements. While qualitative (or semiquantitative) comparisons may be useful, more information potentially may be derived from quantitatively modeling the responses. As a first step towards addressing the issues associated with evaluating dose-response concordance, the following approach was implemented.

Following standard EPA practice (U.S. EPA, 2005), a multistage model was fit to each tumor data set. The degree of the polynomial in that model was set to one less than the number of dose groups. Note that such a model predicts the probability of tumor response as a function of dose. The EPA software, BMDS (U.S. EPA, 2003), was used to obtain and evaluate the model fits.

For MF, a continuous (or pseudocontinuous) endpoint, a quadratic model was used to relate the mean response to the dose level, i.e., $m(d) = b_0 + b_1*d + b_2*d^2$, where d is dose, m(d) is the mean response (MF) at dose d, and the b_i parameters are to be estimated. It was assumed that the MF observations were normally distributed around those means with a variance independent of dose level. BMDS was used to fit this model to the data for riddelliine and DCA.

The model fit to the MF data predicts mean response as a function of dose. It was considered appropriate, for considering dose-response concordance between the mutation response and the tumor response, to express the modeling results in terms of the probability of an "adverse" response. The idea underlying the conversion of a continuous response to the probability scale is familiar from the definition of benchmark doses for continuous endpoints derived from the so-called hybrid approach (Kodell and West, 1993). It is equivalent to specifying some background probability of adverse response (the definition of which is discussed below), thus implicitly

identifying a cutoff value for an adverse response. Then, using the model-estimated standard deviation and dose-related change in the means, the increased probability of having an adverse response is calculated as a function of dose.

Model Fitting

The data were fit using BMDS, free software from the U.S. EPA (U.S. EPA, 2003). Summaries of the fitted model estimates are given here. Figure 2 shows the fit of the models to the data.

Riddelliine Cancer Model: $P(d) = 1 - \exp(- 0.2405*dose^2 - 1.1958*dose^3)$.

Goodness of fit p-value: 0.8124

Riddelliine Mutant Frequency Model: $m(d) = 33.2521 + 84.7974*dose - 15.2044*dose^2$; the constant variance, v(d), was estimated to be 140 for all d.

Goodness of fit p-value: 0.1539

DCA Cancer Model: $P(d) = 0.2772 + (0.7228)[1 - \exp(- 0.6147*dose - 0.2017*dose^2 - 0.0844*dose^3 - 0.00008*dose^5)]$.

Goodness of fit p-value: 0.5789

DCA Mutant Frequency Model: $m(d) = 41.10 + 10.8057*dose + 1.3943*dose^2$; the constant variance, v(d), was estimated to be 115 for all d.

Goodness of fit p-value: NA (same number of parameters as independent mean model; model predictions were the same as that model so the fit to observed means was perfect).

Calculation of Probability of Response for Mutant Frequency

The conversion of the modeled dose-response for the mean (and variance) of MF as a function of dose was accomplished using the following equation:

$$P(d) = 1 - N(m(d), v(d), m(0) + v(0)*N^{-1}(0, 1, 1-b))$$

where N(x, y, z) is the probability that a normal random variable with mean x and variance y will be less than z; $N^{-1}(r, s, t)$ is the inverse normal function giving the value such that the cumulative probability that a normal random variable with mean r and variance s will be equal to t; and b is the assumed background probability of response. The values of m(d) and v(d) are as determined by the fit of the dose-response models to the MF data, shown above.

A key consideration in conducting the modeling for the MF data was the definition of the background response rate. A background response rate of 0.01 is often implicitly assumed in benchmark dose analyses for continuous endpoints (U.S. EPA, 2000). However, it is not clear that this is appropriate for the current application, because the question being addressed is not what level of the MF per se is adverse, but what level would lead to an adverse response (increase) in tumor incidence. Therefore, in addition to the 0.01 background response, the implications of using alternative background responses were considered.

RESULTS

Time-to-Mutation

For DCA, there was no increase in the MF at either 4 or 10 weeks, but there was an increase in MF at 60 weeks (Leavitt *et al.*, 1997). This provides evidence that the time-to-mutation for DCA may be very long (Figure 1a). Thus, the time-to-mutation curve is consistent with DCA being a nonmutagenic carcinogen, since an increase in MF was detected only following long-term exposure. As shown in Figure 1b, riddelliine induces a clear positive response at the earliest (and only) evaluated time point (12 weeks of exposure) while DCA is nonmutagenic after a comparable time period (10 weeks of exposure). Although different units are used on the x-axis for the two chemicals, it is appropriate to compare the curves, which show the mutagenic response in the tumorigenic dose range.

Dose-Response Concordance

Riddelline

Qualitative comparison of the dose-response data (Table 1 and Figure 2a) indicates that liver MF (at 12 weeks of treatment) leads liver tumor incidence. Of the doses tested (0.1, 0.3, and 1.0 mg/kg/day), a statistically significant increase in tumors was observed only at 1.0 mg/kg/day, while all three doses produced statistically significant dose-related increases in MF. Results of the dose-response modeling for the two endpoints are shown in Figure 2a. The calculated probabilities of an adverse response are shown in Figure 3a. The two curves shown for the probabilities associated with MF correspond to two choices of the background rate of response, either 0.058 (the 95% upper bound on the observed rate of response of 0 out of 50 observed in the cancer control group) or 0.01.

With either assumed background response rate for MF, the curve for mutation "leads" the curve for cancer. At any given dose, the likelihood of observing an adversely high MF is greater than the likelihood of detecting a tumorigenic response. Thus, both the qualitative analysis and the modeling results are consistent with a mutagenic MOA for tumor induction.

Dichloroacetic Acid

Complete qualitative comparison of the DCA data is difficult because MF measurements were not made on doses less than 1.0 mg/ml. Sixty weeks of dosing with 1.0 or 3.5 mg/ml of DCA in drinking water produced significant increases in liver MF, while doses of 0.5 to 3.5 mg/ml of DCA produced dose-related increases in liver tumor incidence (Table 2 and Figure 2b).
Plots of the dose-response data and model predictions based on maximum likelihood estimation are given in Figure 2b for both endpoints. The calculated probabilities of an adverse response are shown in Figure 3b. Again, two curves are shown for MF, differing because of different assumptions about the background rate. In one case, the background probability of response was set equal to the background incidence of the tumors in the control group (0.26); in the other case, the background rate is the "default" choice of 0.01.

When the background rate of response for MF is assumed to be small (0.01), the dose-response for cancer leads the dose-response for mutation. When the background rate of response for MF is assumed to be equal to that of the tumor endpoint, the mutant and cancer dose-response curves are essentially indistinguishable. Even though the analysis for DCA appears less clear than that for riddelliine, it still can be concluded that the dose-response concordance for DCA is different from that for riddelliine and is more consistent with a nonmutagenic MOA. In support of this

conclusion, the observed increase in mutations occurred relatively late in the exposure; 10 weeks of treatment with 3.5 mg/ml DCA did not significantly increase liver MF (Figure 1b).

DISCUSSION

Based on an understanding of tumor etiology, we propose a new experimental approach for using mutation data to inform cancer MOA, with particular attention to evaluating whether the dose-response and temporality for the gene mutation data are consistent with mutation being a key event for the formation of tumors. While research is needed to define an optimal design for the experimentation and analysis, we can identify key components of the approach. The species, strain, route of exposure, dose range, and the tissues evaluated for mutation will be selected based on the cancer bioassay. A sufficient number of doses (based on the cancer bioassay) will be employed to provide an assessment of both the time-to-mutation and the dose-response. It is important to include enough doses to provide confidence in dose-response modeling and, in particular, to include one or more doses below those required to induce tumors. The time-to-primary DNA damage, time to mutation, and time to other relevant biological effects (temporality) and dose response concordance information can be used as outlined in the EPA Cancer Guidelines (U.S. EPA, 2005) and the new EPA Framework Document (http://www.epa.gov/osa/mmoaframework/pdfs/MMOA-ERD-FINAL-83007.pdf.) as a part of the WOE evaluation for the MOA Framework. In light of the complexity of these assays and the time and expense involved in conducting these experiments, we envision that the approach described here would be of interest only for chemicals that have been identified as mutagens in short-term tests and as carcinogens, but for which there is controversy regarding the chemical's MOA. In such cases, the approach improves the scientific basis for establishing an MOA.

There are some issues related to these comparisons that are specific to evaluations of gene mutation data. In particular, the mutations being measured are markers for the tumors, but are not on the direct pathway to the tumors, since they are in a reporter gene. However, as noted above, these reporter genes provide a reasonable approximation of the MF in the target gene(s), and, in particular, measure the types of mutations likely to be involved in producing linear dose-responses. We concede that obtaining time-to-mutation and dose-response concordance information using the appropriate oncogenes/tumor suppressor genes would remove a degree of uncertainty from the assessment, but, from a technical standpoint, such analyses are extremely challenging (McKinzie et al., 2001). We also concede that our approach does not readily distinguish between chemicals that are directly DNA reactive and those that are not "directly DNA reactive" yet rapidly induce mutations, perhaps via an oxidative MOA for mutation. By incorporating measures of oxidative damage and assessing time to oxidative damage and time to mutation induction and the dose-response concordance of both endpoints, we feel that we can provide information that contributes to the WOE assessment as to whether the induction of mutation is "direct" or "indirect". The incorporation of antioxidants into the experimental design may assist with this assessment. In spite of these issues, we believe that our approach provides a practical means for providing MOA assessments that are far more scientifically rigorous than the current WOE evaluation of standard genetic toxicology assay data to determine whether the chemical has a mutagenic MOA for the induction of tumors.

Other issues relevant to the dose-response comparisons apply in general to the evaluation of potential key events that are continuous endpoints. To date, comparisons of dose-response for

continuous endpoints typically have used a NOAEL/LOAEL approach (e.g., for chloroform, U.S. EPA, 2006). However, in light of the well-known limitations to the NOAEL/LOAEL method (Crump, 1984), it would be useful to bring a quantitative analytical approach to the evaluation of dose-response concordance between potential key events and tumors. Determination of an appropriate approach for identifying the number of animals with the background level of adverse response, and how to define that background level, is an important consideration for the quantitative analyses of key events described by continuous data. Little research exists to inform this decision. Our examples show that background probability assumptions in the modeling of a continuous endpoint can affect the appearance of dose-response concordance.

In order to begin to evaluate the impact of assumptions regarding the background degree of adversity in MF, two approaches were used. The first approach employed a background rate of 0.01, as is done commonly for continuous endpoints. A second approach used the background tumor rate (or an estimate thereof). This approach was chosen to tie together the MF and tumor data, because the definition of adversely high MF is the MF that increases the risk of tumors. If the mutation data were indicative of a mutagenic MOA, then the implicit rate of response (probability of "adversely" high MFs) should correspond to the rate of the induced tumor response. Equating the background rate of response may be a feasible approach, and it avoids any default choice made in the absence of any knowledge of what values of a continuous endpoint might be considered adverse.

Even with this as a guide, however, there are still other issues. In the riddelliine example, the observed tumor rate in the controls was 0. However, it is not possible to assign 0 background probability of adversely high MFs. The background probability of response implicitly defines a cutpoint above which the MF is "adverse." For any finite cutpoint, there will be some nonzero (albeit perhaps small) probability of an adverse response, meaning that the background probability must be nonzero. In our analyses of riddelliine, we examined the impact of setting the background to some upper bound based on the concurrent control sample size. (Our choice, 0.058, is the 95% upper bound on the binomial probability when 0 out of 50 animals were observed to have the tumor of interest). Examination of historical control data also might be useful for constraining the assumed background rate.

We also have opted not to present statistical analyses of bounds on the probability curves shown here. It may be important to consider more formal ways of approaching the question of whether one curve "leads" another. Comparison of appropriately chosen upper and lower bounds on such curves is one approach that we have begun to explore. Moreover, approaches that do not restrict attention to the tails of the continuous MF distributions and that integrate MF data and tumor response data in a likelihood estimation approach may be useful. These ideas are important for fixing notions of dose-response concordance, not only for mutagenicity data and MOA determination, but whenever continuous and dichotomous responses are being compared. We are continuing to explore the implications of these alternative approaches to evaluating dose-response concordance in ongoing work.

Another important issue, not addressed in our current modeling, is the fact that some chemicals likely have a mixed MOA. That is, they may induce cancer via both mutagenic and nonmutagenic MOAs. To further complicate the modeling analysis, the degree to which these two MOAs contribute may be dependent upon dose. For example, the mutagenic MOA may dominate with a low slope at low doses, with the nonmutagenic MOA contributing only at higher doses, but having a steeper slope. We currently are considering these issues and are developing an additional strategy for dealing with this type of data.

In conclusion, although the ideal in vivo mutation data sets are not available at present, we have conducted some preliminary analyses with existing data from experiments designed for hazard identification. From these analyses, we conclude that our approach for using in vivo mutation data to inform cancer MOA assessment has merit and we plan to proceed with defining optimal studies and the methodology to analyze them. While qualitative or semiquantitative comparisons of the data may suffice for some applications, we feel that analytical modeling will further aid in informing the cancer MOA evaluation. Enhancements of dose-response approaches for mutation data, as well as for tumor data, should be used to extend the more qualitative assessments.

REFERENCES

Chan, P.C., Haseman, J.K., Prejean, J.D., Nyska, A., 2003. Toxicology and carcinogenicity of riddelliine in rats and mice. Toxicol. Lett. 144**,** 295-311.

Chen, T. Mittelstaedt, R.A., Aidoo, A. Hamilton, L.P., Beland, F.A., Casciano, D.A., Heflich, R.H., 2001. Comparison of hprt and lacI mutant frequency with DNA adduct formation in N-hydroxy-2-acetylaminofluorene-treated Big Blue® rats. Environ. Mol. Mutagenesis 37, 195-202.

Cimino, M.C., 2006. Comparative overview of current international strategies and guidelines for genetic toxicology testing for regulatory purposes. Environ. Mol. Mutagenesis 47, 362-390.

Clewell, H.J., Andersen, M.E., 2004. Applying mode-of-action and pharmacokinetic considerations in contemporary cancer risk assessments: An example with trichloroethylene. Crit. Rev. Toxicol. 34(5):385-445.

Cohen, S.M., Meek, M.E., Klaunig, J.E., Patton, D.E., Fenner-Crisp, P.A., 2003. The human relevance of information on carcinogenic modes of action: Overview. Critical Rev. in Tox. 33, 581-589.

Crump, K., 1984. A new method for determining allowable daily intakes. Fundam. Appl. Toxicol. 4, 854-871.

DeAngelo, A.B., George, M.H., House, D.E., 1999. Hepatocarcinogenicity in the male B6C3F1 mouse following a lifetime exposure to dichloroacetic acid in the drinking water: dose-response determination and modes of action. J. Toxicology and Environmental Health, Part A 58, 485-507.

Dearfield, K.L., Moore, M.M., 2005. Use of genetic toxicology information for risk assessment. Environ. Mol. Mutagenesis 46**,** 236-245.

Dearfield, K.L., Auletta, A.E., Cimino, M.C, Moore, M.M., 1991. Considerations in the U.S. Environmental Protection Agency's testing approach for mutagenicity. Mutation Res. 258**,** 259-283.

Dycaico, M.J., Provost, G.S., Kretz, P.L., Ransom, S.L., Moores, J.C., Short, J.M., 1994. The use of shuttle vectors for mutation analysis in transgenic mice and rats. Mutation Res. 307**,** 461-478.

Fu, P.P., Siz, Q., Lin, G., Chou, M.W., 2004. Pyrrolizidone alkaloids – genotoxicity, metabolism enzymes, metabolic activation, and mechanisms. Drug Metab. Rev. 36, 1-55.

Hanahan, D., Weinberg, R.A., 2000. The hallmarks of cancer. Cell 100**,** 57-70.

Hayashi, H., Kondo, H., Masumura, K., Shindo, Y., Nohmi, T., 2003. Novel transgenic rat for in vivo genotoxicity assays using 6-thioguanine and Spi- selection. Environ. Mol. Mutagenesis 41**,** 253-259.

Heddle, J.A., Dean, L., Nohmi, T., Boerrigter, M., Casciano, D., Douglas, G.R., Glickman, B.W., Gorelick, N.J., Mirsalis, J.C. Martus, H.-J., Skopek, T.R., Thybaud, V., Tindall, K.R., Yajima, N., 2000. In vivo transgenic mutation assays. Environ. Mol. Mutagenesis 35**,** 253-259.

Heddle, J.A., Martus, H.J., Douglas, G.R., 2003. Treatment and sampling protocols for transgenic mutation assays. Environ. Mol. Mutagenesis 41**,** 1-6.

Hill, K.A., Buettner, V.L., Glickman, B.W., Sommer, S.S.,1999. Spontaneous mutations in the Big Blue transgenic system are primarily mouse derived. Mutat. Res. 436, 11-19.

Kavlock, R.J., Schmid, J.E., Setzer, R.W., Jr., 1996. A simulation study of the influence of study design on the estimation of benchmark doses for developmental toxicity. Risk Analysis 16, 399-410.

Kirman, C.R., Gargas, M.L., Marsh, G.M., Strother, D.E., Klaunig, J.E., Collins, J.J., Deskin, R., 2005. Cancer dose--response assessment for acrylonitrile based upon rodent brain tumor incidence: use of epidemiologic, mechanistic, and pharmacokinetic support for nonlinearity. Regul. Toxicol. Pharmacol. 43(1):85-103.

Knudson, A.G., 2002. Cancer genetics. Am. J. Med. Genet. 22, 96-102.

Kodell, R.L. West, R.W., 1993. Upper confidence limits on excess risk for quantitative responses. Risk Analysis 13**,** 177-182.

Leavitt, S.A., DeAngelo, A.B., George, M.H., Ross, J.A., 1997. Assessment of the mutagenicity of dichloroacetic acid in lacI transgenic B6C3F1 mouse liver. Carcinogenesis 18**,** 2101-2106.

McKinzie, P.B., Delongchamp, R.R., Heflich, R.H., Parsons, B.L. 2001. Prospects for applying genotypic selection of somatic oncomutation to chemical risk assessment. Mutat. Res. 489, 47-78.

Meek, M.E., Bucher, J.R., Cohen, S.M., Dellarco, V, Hill, R.N., Lehman-McKeeman, L.D., Longfellow, D.G., Pastoor, T., Seed, J., Patton, D.E., 2003. A framework for human relevance analysis of information on carcinogenic modes of action. Crit. Rev. Toxicol. 33, 591-653.

Mei, N., Heflich, R.H., Chou, M.W., Chen, T., 2004a. Mutations induced by the carcinogenic pyrrolizidine alkaloid riddelliine in the liver cII gene of transgenic Big Blue rats. Chem. Res. Toxicol. 17, 814-818.

Mei, N., Chou, M.W., Fu, P.P., Heflich, R.H., Chen, T., 2004b. Differential Mutagenicity of riddelliine in liver endothelial and parenchymal cells of transgenic big blue rats. Cancer Lett. 215, 151-158.

Melnick, R.L., Huff, J., Barrett, J.C., Maronpot, R.R., Lucier, G, Portier, C.J., 1993. Cell proliferation and chemical carcinogenesis: a symposium overview. Mol. Carcinogenesis 7, 135-138.

Moore, M.M., Harrington-Brock, K., 2000. Mutagenicity of trichloroethylene and its metabolites: implications for the risk assessment of trichloroethylene. Environ. Health Perspect. 108(Suppl 2), 215-223.

National Toxicology Program, 2003. Toxicology and carcinogenesis studies of riddelliine (CAS No. 23246-96-0) in F344/N rats and B6C3F1 mice (gavage studies). Natl. Toxicol. Program Technol. Rep Ser 508, 1-280.

Okonogi, H., Ushijima, T., Zhang, X.B., Heddle, J.A., Suzuki, T., Sofuni, T., Felton, J.S., Tucker, J.D., Sugimura, T., Nagao, M. 2001. Agreement of mutational characteristics of heterocyclic amines in *lacI* of the Big Blue® mouse with those in tumor related genes in rodents. Carcinogenesis 18, 745-748.

Shipp, A., Lawrence, G., Gentry, R., McDonald, T., Bartow, H., Bounds, J., Macdonald, N., Clewell, H., Allen, B., and Van Landingham, C. (2006). "Acrylamide: Review of toxicity data and dose-response analyses for cancer and noncancer effects." Crit. Rev. Toxicol. 36(6-7), 481-608.

Skopek, T.R., Kort, K.L., Marino, D.R., 1995. Mutation of the endogenous *hprt* gene and *lacI* transgene in ENU-treated Big Blue™ B6C3F1 mice. Environ. Mol. Mutagenesis 26, 9-15.

Thilly, W., 2003. Have environmental mutagens caused oncomutations in people? Nat. Genet. 34(3), 255-259.

Thybaud, V., Dean, S., Nohmi, T., de Boer, J., Douglas, G.R., Glickman, B.W., Gorelick, N.J., Heddle, J.A., Heflich, R.H., Lambert, I., Martus, H.J., Mirsalis, J.C, Suzuki, T., Yajima, N., 2003. In vivo transgenic mutation assays. Mutation Res. 540, 141-151.

U.S. EPA. 2000. Benchmark Dose Technical Guidance Document. External Review Draft. EPA/630/R-00/001, October 2000.

U.S. EPA. 2003. EPA's Benchmark Dose Software (BMDS), Version 1.4.1. Available at: www.epa.gov/ncea/bmds.htm.

U.S. EPA. 2005. Guidelines for Carcinogen Risk Assessment. EPA/630/P-03/001B, March 2005.

U.S. EPA. 2006. U.S. Environmental Protection Agency 'Integrated Risk Information System (IRIS)', Available at http://www.epa.gov/iris/subst/0410.htm.

U.S. EPA, 2007. Framework for Determining a Mutagenic Mode of Action for Carcinogenicity, available at http://www.epa.gov/osa/mmoaframework/pdfs/MMOA-ERD-FINAL-83007.pdf.

Zhang, X.B., Felton, J.S., Tucker, J.D., Urlando, C., Heddle, J.A., 1996. Intestinal mutagenicity of two carcinogenic food mutagens in transgenic mice: 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine and amino(α)carboline. Carcinogenesis 17, 2259-2265.

Table 1. Liver Mutant Frequency (MF) and Hemangiosarcoma Data for Rats Dosed with Riddelliine

Dose (mg/kg/day)[a]	MF data after 12 weeks of dosing[b]		Hemangiosarcoma data after 104 weeks of dosing[c]		
	No. of rats	MF ± SD ($x10^{-6}$)[d]	No. of rats examined	No. of rats with tumors	% with tumors
0	6	30 ± 10	50	0	0
0.1	6	47 ± 14*	50	0	0
0.3	6	55 ± 8**	50	3	6
1.0	6	103 ± 16***	50	38	76†

Significantly greater than control (Tukey test): *p<0.05, **p<0.01, ***p<0.001.

Significantly greater than control (Poly-K test): †p<0.05.

[a]Rats were dosed 5 days per week with the indicated dose.

[b]Data from Mei *et al.* (2004a).

[c]Data from the National Toxicology Program (2003).

[d]MF is expressed as the number of mutants per 10^6 phages. Assays were conducted on transgenes recovered from whole liver, with each liver cell containing approximately 20 copies of the transgene.

Table 2. Liver Mutant Frequency (MF) and Hepatocellular Carcinoma Data for Mice Dosed with Dichloroacetic Acid

Dose (mg/ml)	MF data after 60 weeks of dosing[a]		Hepatocellular carcinoma data after 100 weeks of dosing[b]		
	No. of mice	MF ± SD ($x10^{-6}$)[c]	No. of mice examined	No. of mice with tumors	% with tumors
0	5	41 ± 10	50	13	26
0.05	ND	ND	33	11	33
0.5	ND	ND	25	12	48†
1.0	5	53 ± 10*	35	25	71††
2.0	ND	ND	20	19	95††
3.5	6	96 ± 15**	11	11	100††

ND, not done

Significantly greater than control (Cochran-Armitage test): *p=0.05, **p=0.01.

Significantly greater than the control (Fisher's exact test): †p=0.05, ††p≤0.05

[a]Data from Leavitt *et al.* (1997).

[b]We thank Dr. Anthony DeAngelo for providing the primary liver carcinoma data from his research published in DeAngelo *et al.* (1999).

[c]MF is expressed as the number of mutants per 10^6 phages. Assays were conducted in transgenes recovered from whole liver, with each liver cell containing approximately 40 copies of the transgene.

FIGURE LEGENDS

Figure 1. (A) Time-to-mutation for animals exposed to DCA. The induced MF (the observed MF less the concurrent background MF) is plotted. MF was evaluated at 4, 10, and 60 weeks of exposure. Only the 60-week exposures of DCA were positive. **(B)** The induced MF (the observed MF less the concurrent background MF) for animals exposed to either riddelliine or DCA for either 12 or 10 weeks, respectively, is shown to emphasize the difference in the riddelliine and DCA mutant induction responses.

Figure 2. The observed MF and cancer incidence are plotted with the model predicted mutant frequency and cancer incidence. **(A)** Results for riddelliine, and **(B)** results for DCA.

Figure 3. (A) Fitted dose-response curves for MF and cancer incidence on a probability scale for riddelliine. Cancer dose-response (————) is based on fitting a multistage model. Two probability dose-responses for MF are shown, resulting from different assumptions regarding the background rate of response (0.058, — — —; or 0.01, -------). **(B)** Fitted dose-response curves for MF and cancer incidence on a probability scale for DCA. Cancer dose-response (————) is based on fitting a multistage model. Two probability dose-responses for MF are shown, resulting from different assumptions regarding the background rate of response (0.26, — — —; or 0.01, -------).