UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

OFFICE OF
PREVENTION, PESTICIDES AND
TOXIC SUBSTANCES

June 5, 2006

MEMORANDUM

SUBJECT: Review of the TERA Document: "Use of Benchmark Concentration Modeling and Categorical Regression to Evaluate the Effects of Acute Exposure to Chloropicrin Vapor". MRID 46614801

DP Barcode D305307
TXR No.: 005422
PC Codes: 081501

FROM: Elissa Reaves, Ph.D., Toxicologist/Risk Assessor
Reregistration Branch 2
Health Effects Division (7509C)

THROUGH: William Hazel, Branch Chief
Reregistration Branch 2
Health Effects Division (7509C)

TO: Nathan Mottl, Chemical Review Manager
Special Review Branch
Special Review and Reregistration Division (7508C)

This document presents a review of the modeling performed by the Toxicology Excellence for Risk Assessment (TERA) group (MRID 46614801). Upon the request by the Chloropicrin Manufacturers Task Force, the TERA group analyzed data from the chloropicrin human study (MRID 46443801) for possible use in the human health risk assessment for chloropicrin.

cc: Tina Levine
Jack Housenger

1. Background

Benchmark dose (BMD) analysis, or in this case benchmark concentration (BMC) analysis is generally a preferable alternative to the no-observed-adverse-effect level (NOAEL)/lowest-observed-adverse-effect-level (LOAEL) approach to define a point of departure (PoD). This approach has several advantages over the traditional NOAEL/LOAEL approach in that the BMC considers the dose-response of all the data, is not confined to tested exposure concentrations, is not dependent upon a study NOAEL, and more appropriately accounts for study size and thus study power. This approach provides both the central estimate, also referred to as the BMC, and the corresponding confidence limit (BMCL) for the BMC. The Office of Pesticide Programs is increasing its use of BMD/BMC techniques in its human health risk assessments. The BMC and categorical regression analysis by TERA was submitted by the Chloropicrin Manufacturers Task Force. The Agency appreciates the effort and completeness of the modeling approaches undertaken by TERA for the chloropicrin data and encourages the submission of similar data analyses.

TERA used three different software programs in the chloropicrin analysis. EPA's Benchmark Dose Software (BMDS) was used for determining the most appropriate BMC model for ocular irritation data from the single chloropicrin human study. ToxTools (Cytel) software, however, was also used to model the ocular symptom data to control for intra-subject correlations that are not accounted for by the BMDS program. EPA's CatReg modeling program, a customized S-Plus (MathSoft, Inc.) software package, was utilized by TERA for the present categorical regression analyses. Categorical regression allows the combination of dichotomous data (also known as quantal data, i.e., incidence of a histopathology lesion), continuous data (e.g., serum levels of liver enzymes), and descriptive data (e.g., "severe lung lesions were observed at x dose") from one or more studies into a single analysis. In the case of the available chloropicrin human and animal data, use of these three modeling approaches appears reasonable.

The TERA group analyzed data from the chloropicrin human study (MRID 46443801 also referred to as Cain, 2004). Briefly, the chloropicrin human study consists of three phases, each phase varying in duration and exposure concentration. Please refer to the Data Evaluation Record (DER) of the chloropicrin human study for more details (DP312312). Results of the human study revealed that ocular irritation is the most sensitive endpoint, and therefore, is the focus of the analyses provided by TERA. Specifically, TERA focuses on phase 3 ocular irritation data for BMC and categorical regression modeling. Phase 3 of the human study consisted of 60 minutes of chloropicrin vapors for 4 consecutive days. This one hour repeated exposures most closely relates to the acute bystander scenario for the chloropicrin human health risk assessment. All available chloropicrin animal data were included in the categorical regression analysis except for the developmental studies (rat and rabbit) and the 2-generation reproduction study (rat). The design of these studies differed from the

other longer-term studies, so that the observed severities were not directly comparable.

2. Summary of TERA Analyses

The main modeling approaches for the chloropicrin data are outlined below. The overall goal of TERA was to employ several approaches of using the data for BMC (human data only) and categorical regression (both human and animal data) analyses. The chloropicrin human study has a complicated design with the same subjects exposed to multiple concentrations on multiple days which complicate the BMC analyses. The four main methods of analyzing the chloropicrin data are briefly described below.

The first main approach to the BMC modeling involved the use of all response data from the 4 consecutive exposure days and to each concentration. The impact of multiple days of exposure on the BMC analysis is tested by treating the data as either independent (each day is independent event), clustered (account for same individual exposed on multiple days and to each concentration), or semi-clustered (account for same individual exposed on multiple days but not account for each concentration).

i. Averaged ocular response data across days (4 days total):

a. Independent approach using BMDS
b. Clustered approach using ToxTools software
c. Semi-clustered approach using BMDS

The second method of analyzing the ocular data was to average the response data across the middle of the plateau period (30-55 minutes) that was evident from the human study. The goal of this method was to capture the period of maximal response. An alternative to this approach would be to average across the entire 1-hour period of exposure. Averaging across the entire period, however, would underestimate the response, and result in a higher BMC. This approach was not explored further, since it was not considered representative of the exposure duration of interest.

ii. Averaged response ocular data at plateau period of phase 3 (30-55 minutes)

The third method evaluated the implication of choice of ocular irritation score as the cutpoint for determining a positive response. This alternative method used averaged response data (during the plateau period) but only as a response incidence based on different ocular irritation scores from the human study.

iii. Average response ocular data at plateau period but incidence of positive response based on various ocular severity cutpoints (i.e., 1, 1.5, 2, 2.5)

The fourth method involved the categorization of the animal and human data by severity on an individual level (i.e., based on the judgment of severity for each subject or test animal) for categorical regression analysis. The challenge in this method was to be consistent in defining a severity level across animal and human data, regardless of the endpoint. All modeling was conducted based on the exposure concentration, without additional adjustments for human equivalent concentrations (HECs), which take into account interspecies (animal to human) differences in tissue dose. To avoid apparent inconsistencies in the concentration-duration-response from different HECs for the same concentration at different durations, the initial analysis did not include the HECs.

iv. Averaged ocular response data across 4 days for EC_{10} lower bound value (categorical regression):

a. Data categorized by severity on an individual level
b. Severity levels based on 4 categories
c. Categorization of animal data without HEC (human equivalent concentration) adjustments
d. Round average values down or up.

Table 1 outlines the choice of model used, the alternative approach, and the impact of the modeling (i.e., point of departure) as provided in the TERA report. Lower bound estimates from both the BMC and categorical regression analyses range approximately from 70 ppb to 90 ppb. The TERA suggest a 1-hour exposure limit be derived from the $BMCL_{10}$ of 73 ppb for ocular symptoms.

3. BMC Modeling of Human Data

The uncertainties surrounding the modeling approaches for the phase 3 data of the chloropicrin human study consist mainly of continuous data, repeated measures, and possible carry-over effects that may not be accommodated in the BMC model. The likely impact of these uncertainties on the resulting BMCL is addressed below.

a. Nature of data: continuous & quantal

First, the most appropriate model for typical continuous data (e.g., serum levels of liver enzymes) would include linear and polynomial models, power models or other nonlinear models such as Hill models. However, the self-reported irritation scores (scores of 0, 1, 2, or 3) from phase 3 of the chloropicrin human study are not typical continuous data. In this instance, therefore, the use of the irritation scores to model as a quantal value is appropriate. Our conclusion, therefore, is that the TERA analysis is adequate for determining a BMCL based on continuous subject irritation data.

b. Repeated measures

Phase 3 of the human study is a repeated measures design since subjects were exposed repeatedly throughout the experiment. The data should therefore be analyzed with repeated measures techniques. The BMDS model, however, typically does not consider repeated measures. The ToxTools software, also employed in the analysis, allowed subjects to be clustered. This approach accounts for the same individual being exposed on multiple days and to each concentration. ToxTools also allows dose to differ among subjects in a cluster whereas the nested models in BMDS do not cluster. Although the data were modeled several different ways, which address repeated measures with the data, the range of the lower bound estimates is narrow. The similar lower bound estimates to the different modeling approaches suggest that the impact of the repeated measures in the analyses is small. Our conclusion, then, is that the TERA analysis is adequate for addressing the repeated measures data from the human study.

c. Carry-over effect

The last uncertainty to consider in the TERA analyses is the possible carry-over effect from consecutive exposures. Eye irritation from the first exposure, for example, may still be present when the next consecutive exposure begins. Phase 3 of the human study consisted of 4 consecutive days of exposure. The statistical analysis reported in the human study indicates an ANOVA of Level by Day for ocular irritation of $p \leq 0.02$. However, the human study report also indicates this ANOVA significance held no meaning since there was a lack of progressivity of eye irritation across days. So, our conclusion is that the contribution of any carry-over effect on the TERA analysis, therefore, would be expected to be small.

4. Categorical Regression of Both Animal and Human Data

The key points and/or uncertainties of the categorical regression analyses include intra-individual variability, lack of adjustments for human equivalent concentrations (HECs) of animal data, and scaling of severity level among the data. The likely impact of these uncertainties on the resulting EC_{10} is addressed below.

a. Intra-individual variability

The intra-individual variability, also known as intraspecies variability, was likely not captured in the categorical regression since the ocular symptom data that was generated over 4 replicate days was averaged for the analysis. The focus of this analysis, however, was not on variability of subject response but to inform the impact of chloropicrin concentration and duration beyond the 1-hour human data.

b. No HEC (human equivalent concentration) conversion

The second key point of the categorical regression is the lack of conversion of animal data to HECs. For acute respiratory effects to vapors, the EPA guidance indicates the animal HEC is the same as the human exposure concentration. The optimal comparison, therefore, would be based on HECs. TERA indicated, however, that HECs were not included to avoid apparent inconsistencies in the concentration-duration-response from different HECs for the same concentration at different durations. It may be concluded, therefore, that the lack of HEC conversion of the animal exposure concentrations may potentially underestimate the chloropicrin concentration estimated at 8 and 24 hours.

c. Categorization of severity level

A third key aspect of the categorical regression analyses is the common definition of severity level, regardless of the endpoint. The approach taken by TERA has the effect of equating overt toxicity observed in animal studies (histopathological changes) with self-reported signs in human subjects where chloropicrin may be detected but with no clinical affects. The conclusion that may be drawn from this approach is that best professional judgment is required for assignment of severity levels in a consistent method.

d. Fit of data into categorical regression models

Finally the fit of the animal and human data into the categorical regression models was not as good as the fit into the BMC models. The likely explanation for a poorer fit than the BMC modeling is the fact that only one set of data from one study was used in the BMC analyses, whereas the categorical regression is a meta-analysis of multiple studies, and attempts to fit all the data. The fit of the data based on four severity levels, however, appears consistent with the human data and the biology. So, it may be concluded that the overall fit of the animal and human data into the categorical regression models has little impact on the resulting estimated EC_{10} generated in the TERA analysis.

e. Strengths of the Categorical Regression Analysis

The overwhelming strength or aspect of the categorical regression analysis is the extrapolation or estimation of a response to a specified concentration and duration when no data are available. The TERA report estimated chloropicrin concentrations at longer acute exposures (e.g. 8 and 24 hours) to chloropicrin vapor, which are currently not available for either humans or animals. This modeling approach is capable of combining dichotomous data (also known as quantal data, i.e., incidence of a histopathology lesion) with continuous data (e.g., serum levels of liver enzymes) and/or descriptive response data (e.g., “severe lung lesions were observed at x dose”) from one or more studies into a single analysis. The categorical regression, therefore, is an important approach

since the toxicity to chloropicrin is dependent upon both exposure concentration and duration. Our conclusion is that the categorical regression analysis does provide pertinent information for extrapolation of chloropicrin to the 8 hour or 24 hour period.

4. Discussion

Based on the previous discussion, the impact of the uncertainties or issues with the human data on the $BMCL_{10}$ and EC_{10} estimates provided in the TERA report are likely small. These assumptions are supported by the various approaches to modeling the data that resulted in consistent lower bound estimates ($BMCL_{10}$ and EC_{10}). Lower bound estimates, for example, from both the BMC and categorical regression analyses range approximately from 70 ppb to 90 ppb. The TERA suggest a 1-hour exposure limit be derived from the $BMCL_{10}$ of 73 ppb for ocular symptoms. The BMC and categorical regression modeling approaches provided in the TERA report, therefore, appear biologically and statistically robust for estimating a point of departure for the acute bystander scenario of the chloropicrin human health risk assessment. The $BMCL_{10}$ of 73 ppb is based on eye irritation scores from phase 3 of the human study and is appropriate for establishing a PoD for the acute inhalation exposure scenario of the human health risk assessment.

Table 1. Impact of Alternative Choices for Hazard Endpoint Identification and for Assessment Methods (taken from TERA report, page 12 of 314)

Choice Made	Alternative Approaches	Impact (point of departure unless otherwise specified)
BMC Modeling for ocular endpoint (unless otherwise specified)		
1. Use of averaged response across days in determining response incidence BMCL=73 ppb Range of results of individual models with acceptable fit, using either the averaged or individual data: BMCL=71-89 ppb	a) Treat each individual day as an independent event in determining incidence	BMCL=82 ppb
	b) Use a clustered approach to the modeling, taking into account that the same individual was exposed on multiple days, and to each concentration-using ToxTools software	BMCL=82 ppb (based on best fitting model) or 89 ppb (based on average of successful models), but not all of the EPA BMDS models that fit the averaged data best could be run successfully with ToxTools
	c) Use a semi-clustered approach, taking into account that the same individual was exposed on multiple days (but not taking into account that each individual was exposed to each concentration)-using the Nested model in the BMDS software	Model apparently failed; no BMC's available
2. Use of the average response across the middle of the plateau period (30-55 minutes) in Phase 3 of the Cain study BMCL=73 ppb	a) Use of entire range	Would increase BMD. Would be appropriate to include the rising part of curve only for derivation of exposure limits for <0.5 hour, in which case the plateau region would be excluded. This

		approach was not pursued because the focus of this assessment was on 1-hour exposures and longer.
	b) Broaden the time period over which the average is calculated. End the period at the end of exposure (1 hour). For the beginning of the period, estimate the time point that results in the same response as that seen at 1 hour.	This choice did not affect incidence, and so did not affect BMC's.
3. Choice of ocular severity cutpoint of ≥1.5 for average response in determining the response incidence for Phase 3. BMCL=73 ppb	a) Cutpoint of ≥ 0.5	BMCL of 7-23 ppb for various models, but this cutpoint is inappropriate, because it is not consistent with the definintion of adverse; the average response is below "minimal awareness"
	b) Cutpoint of ≥1.0	BMCL=33 ppb. Cutpoint is somewhat outside the reasonable range of cutpoints.
	c) Cutpoint of ≥2.0	BMCL= 120 ppb
	d) Cutpoint of ≥2.5	BMCL=140-142 ppb for various models, but this cutpoint does not appear to be consistent with definition of adverse, in light of the reported subjective
4. Use of average of responses during the plateau period of 30-55 minutes to determine the response incidence in Phase 3 BMCL=73 ppb	a) use holistic evaluation described in Section 7.1 to determine the response incidence	BMCL=42-59 ppb
5. Choice of model used to choose the final BMCL-averaged across those with perfect fit BMCL=73 ppb	a) Average across all studies with acceptable fit in region of interest (based on chi squared-residuals)	BMCL=76 ppb
Categorical Regression Modeling		
1. For the individual score for each time point in Phase 3, the average response across the 4 exposure days was determined. Where the avg. was an intermediate value (0.5, 1,5, etc.) the general approach of rounding up was followed. EC10 Lower bound at 1 hour =90 ppb	a) Round intermediate values (0.5, 1.5, 2.5) down.	1-hour EC10 lower bound of 107 ppb
2. Choice of severity ratings as documented in the text. EC10 lower bound at 1-hour=90 ppb	a) Move all serverity ratings in the Cain study down 1 level (except NOEL cannot be moved further down)	EC10 line not consistent with the data points; 1-hour EC10 lower bound of 241 ppb