17th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
9-12 May 2016, Budapest, Hungary

ARE TOXIC LOAD-BASED TOXICITY MODELS CONSISTENT WITH EXPERIMENTAL OBSERVATIONS? INDEPENDENT ANALYSIS OF STEADY-EXPOSURE DATA FROM THE 2012–2013 ECBC/NAMRU-D TOXICOLOGICAL EXPERIMENTS

Alexander J. Slawik[1], Kevin C. Axelrod[1,2], James B. Silva[1,3], Ivo K. Dimitrov[1], Jeffry T. Urban[1], and Nathan Platt[1]

[1]Institute for Defense Analyses, Alexandria, Virginia, USA
[2]Graduate Program in Biophysics, Harvard University, Cambridge, Massachusetts, USA
[3]Department of Physics, Boston University, Boston, Massachusetts, USA

Abstract: Toxic industrial chemicals and chemical warfare agents present an acute inhalation hazard to civilians and military personnel. An individual exposure to an airborne hazardous material may be highly time-dependent due to the random effects of wind meandering and atmospheric turbulence. Several toxicological models based on the "toxic load" model of exposure have been proposed to predict the casualties arising from time-dependent exposures to airborne hazardous materials, but none were developed using data from toxicological experiments that used time-varying exposure profiles. The US Defense Threat Reduction Agency (DTRA) sponsored a two-year set of experiments, conducted in 2012 and 2013, that were designed and executed through a collaboration between the US Army's Edgewood Chemical and Biological Center (ECBC) and the Naval Medical Research Unit Dayton (NAMRU-D) to explore the effects of time-varying inhalation exposures of hydrogen cyanide (HCN) gas on rats.

Our independent analysis of the data from the ECBC/NAMRU-D toxicological experiments has two components. The first component, which is the subject of this paper, is an analysis of the steady-exposure (square-pulse) data. Since the toxic load model was originally designed with these types of exposures in mind, this analysis is intended to examine whether the basic toxic load model suitably describes the ECBC/NAMRU-D data set. The second component of our analysis, which is described in a separate paper, examines whether several proposed extensions of the basic toxic load model suitably describe the time-varying exposure data from the ECBC/NAMRU-D experiments.

Our analysis of the ECBC/NAMRU-D data indicates that the basic toxic load model is not suitable for describing the steady-exposure data over the full range of the experiments' exposure durations (2.3 minutes to 30 minutes): the model fits the data well only if the short-duration exposures (less than 10 minutes) are dropped from the data set. This is potentially significant since HCN inhalation exposures can be lethal below 10 minutes, so there is a need for toxicological models that can describe toxicity on timescales of not only tens of minutes, but also minutes. Our analysis, however, was unable to determine whether the poor fit was due to systematic experimental error or a failure of the toxic load model to account for the physiological response of the rat across all investigated timescales.

Key words: *Toxicology, toxic load, hazardous materials, acute inhalation exposure, consequence assessment, health and exposure assessment, hydrogen cyanide, HCN, rats*

INTRODUCTION

Health assessments based on atmospheric dispersion modelling require a toxicological model to relate the exposures to airborne contaminants predicted by the dispersion models to the likelihood of adverse health effects arising from those exposures. Our previous work examined whether atmospheric dispersion models that predict ensemble-mean plumes can predict casualties and the location of health hazards accurately as well as some of the theoretical differences between different toxic load-based toxicity models (Urban et al., 2011; Urban et al., 2014; Platt et al., 2014). Our present work examines the experimental validation of toxic load-based toxicity models themselves.

Our research focuses on the health effects arising from accidental or intentional releases of toxic industrial chemicals or chemical warfare agents. For these types of scenarios, acute inhalation exposures

may be on the order of tens of minutes or less if the wind transports the hazardous material beyond the affected population quickly enough, or if the affected population is able to take timely action to mitigate the hazard, such as evacuating from the exposed area, seeking shelter, or donning protective equipment. Acute exposures may be longer, on the order of hours, if the hazard is persistent (e.g., a dense gas in low-wind conditions) and effective hazard mitigation is not possible.

The most common statistical model used in the study of acute inhalation toxicology is the dose-probit response model (Finney, 1947; Sommerville et al., 2006). In this model, toxic effects are quantal (dichotomous), and the likelihood of an organism reaching a given toxic endpoint (e.g., death) depends only on the total amount of toxic material (i.e., the total dose of toxicant) that has accumulated in the organism. The dose-probit model, which is purely phenomenological, does not account for variations in the manner in which the dose is administered, such as the time history of the exposure: all combinations of C and T that have the same product D are predicted to have the same toxic effect. The likelihood of a single organism responding to a given exposure (e.g., by dying) is usually equated with the fractional response of a population of identical organisms subjected to identical exposures – the latter quantity can be measured directly in toxicological experiments. The dose-probit model assumes that the fraction of the population responding, when measured on a probit scale, is a lognormal function of the dosage (Equation 1).

$$Y_P = c_0 + m \log_{10}(D) \tag{1}$$

In this equation, the response function Y_P is measured on a probit scale, m is a quantity called the probit slope, $\log_{10}(z)$ is the base-10 logarithm of z, and the dosage D is assumed to be proportional to the actual inhaled dose, where C is the atmospheric concentration of toxicant in the exposure and T is the exposure duration. The coefficients c_0 and m depend on the type of toxicant and type of organism. This formulation is valid only for steady exposures where C is constant over the whole time interval T. The probit scale linearizes the response function. The actual response function is a sigmoidal function of the logarithm of the dosage D (Eq. 2).

$$P = \left\{\text{erf}\left[\left(m\log_{10}(D) - m\log_{10}(D_{50})\right)/\sqrt{2}\right] + 1\right\}/2 \tag{2}$$

Here P is the fraction of the population that exhibits a given response, erf(z) represents the error function of z, and a new constant has been introduced: D_{50}, the median effective dosage (i.e.., the dosage required to cause a response in 50% of the population). Eqs. 1 and 2 were developed and validated for the case in which C, the atmospheric concentration of toxicant in the exposure, is constant over the whole exposure duration T. In this case the dosage is given by Eq. 3.

$$D = C \times T \tag{3}$$

For some toxicants, particularly some of the toxic industrial chemicals (TICs), the dose-probit model yields a poor fit to data from toxicological experiments. Since the dose-probit model does not take into account the time dependence of physiological processes, such as the uptake and clearance of toxicants in the body, a more general phenomenological model has been proposed to account for the fact that a high-intensity but short exposure could have a different toxic effect from a low-intensity but long exposure of equal dosage. This model, called the "toxic load" probit-response model, treats the concentration-dependence and exposure duration-dependence of the physiological response independently. The "toxic load" generalization of Eqs. 1 through 3 is given by Equations 4 through 6 (ten Berge et al., 1986; Sommerville et al., 2006).

$$Y_P = k_0 + k_1\log_{10}(C) + k_2\log_{10}(T) \tag{4}$$

$$P = \left\{\text{erf}\left[\left(m\log_{10}(TL) - m\log_{10}(TL_{50})\right)/\sqrt{2}\right] + 1\right\}/2 \tag{5}$$

$$TL = C^n \times T \tag{6}$$

In the equations the prior measure of exposure, the dosage D, has been generalized to the "toxic load" TL, which is no longer proportional to the inhaled dose. The coefficients k_0, k_1, and k_2 in Eq. 4, which depend on the type of toxicant and organism, have been rearranged in Eqs. 5 and 6 as the probit slope m, the median lethal toxic load TL_{50}, and the toxic load exponent n.

The toxic load-probit model (Eqs. 4–6), like the dose-probit model, was formulated only for the case of steady-exposures, which are the exposures that have been accessible to toxicological experiments. Real-world exposures to airborne toxicants, however, may be highly time-dependent due to the random effects of wind meandering and atmospheric turbulence (Wilson, 1995). Various extensions to the toxic load model have been proposed to predict the toxicological effects of time-varying exposures $C(t)$, but until recently, little data were available with which to validate these proposed extensions. The toxicological experiments that are the subject of this paper were intended to remedy this deficiency. This paper, however, focuses on the validity of the basic toxic load model rather than its time-dependent extensions.

ECBC/NAMRU-D TOXICOLOGICAL EXPERIMENTS

The US Defense Threat Reduction Agency (DTRA) sponsored a two-year set of experiments, conducted in 2012 and 2013, that were designed and executed through a collaboration between the US Army's Edgewood Chemical and Biological Center (ECBC) and the Naval Medical Research Unit Dayton (NAMRU-D) to explore the toxic effects of time-varying inhalation exposures (Sweeney, et al. 2014; Sweeney, et al., 2015). In these experiments, groups of ten male Sprague-Dawley rats were exposed to HCN gas using a pressurized, nose-only inhalation apparatus that allowed excellent control of the time profile of the exposures. For each exposure, the fraction of the ten rats that did not survive the exposure was recorded.

The ECBC/NAMRU-D experiments investigated three basic types of exposure profile: single steady pulses (i.e., fixed C and T), two unequal-intensity pulses back-to-back, and two unequal-intensity pulses separated by a gap during which there was no exposure to HCN gas. The exposure durations and the ordering of the two-pulse profiles (i.e., high-then-low vs. low-then-high) varied between the 2012 experiments and the 2013 experiments. This paper considers only the steady-pulse data; the two-pulse data are considered in a separate paper. The experiments included 126 trials with ten rats per trial, of which 34 trials featured steady-pulse exposures having various combinations of C and T. In 2012 there were 7 trials with T = 5 minutes, 7 trials with T = 15 minutes, and 6 trials with T = 30 minutes. In 2013 there were 7 trials with T = 2.3 minutes, 5 trials with T = 10 minutes, and 2 trials with T = 30 minutes.

DATA ANALYSIS

We evaluated the basic (i.e., steady-exposure) toxic load model by fitting Eq. 5 with the steady-exposure data from the ECBC/NAMRU-D experiments and then examining measures of the goodness-of-fit. We performed the fits by the maximum-likelihood estimation (MLE) method using the Benchmark Dose Software (BMDS) developed by the US Environmental Protection Agency (EPA, 2015). BMDS estimates the three coefficients of the toxic load model by performing a simultaneous 3-parameter fit. BMDS also estimates the coefficients' covariance matrix using the Fischer information matrix and performs parametric tests of the goodness of fit.

Initial analysis

The BMDS-estimated toxic load model parameters are n = 1.35, $TL_{50} = 5.41 \times 10^4$, and m = 1.68 using all 34 steady-exposure trials. Figure 1 shows the fit of the sigmoidal response function to the 34 data points and a plot of predicted versus observed response using the fitted coefficients. The latter plot is a measure of the self-consistency of the data. A well-fit model would result in predictions that equal observations (i.e., that lie on the diagonal line in a predictions vs. observations plot).

Note that there is considerable scatter in the data. In particular, the 5-minute exposures generally resulted in more casualties than one would expect from the toxic load model, whereas the 2.3-minute exposures generally resulted in fewer casualties than one would expect.

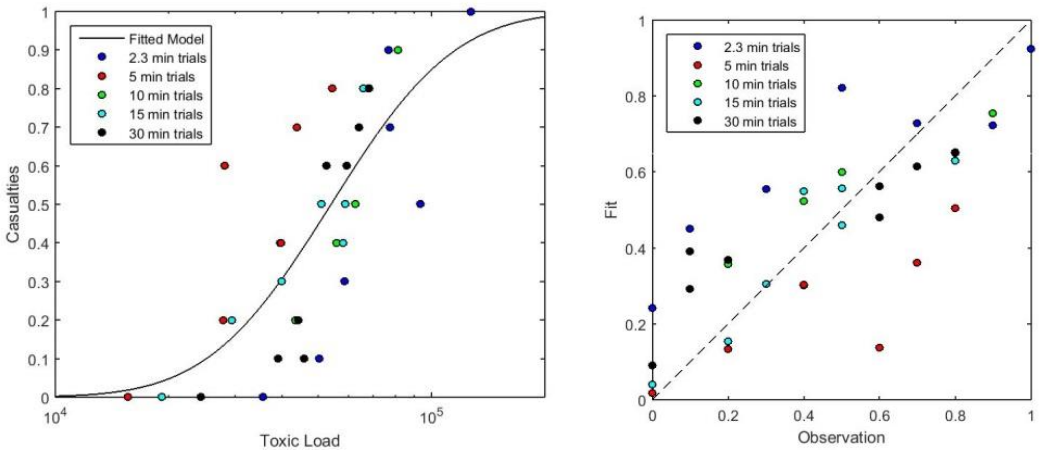

Figure 1. Left: Fit of the toxic load model, Eq. 5, to all 34 steady-exposure trials. Fractional casualties (e.g., 4 rats dead out of 10 = 0.4) are plotted as a function of the toxic load for each trial. Each point represents one trial of 10 rats. Right: Predicted vs. observed fractional casualties for all 34 trials using the coefficients from the fit on the left.

Investigation of the self-consistency of the toxic load model across timescales

Figure 1 indicates that the toxic load model appears not to perform well for the shorter-timescale exposures in the ECBC/NAMRU-D data set. To further explore the performance of the toxic load model at different timescales, we re-fit the model using only subsets of the original 34 trials (e.g., the 20 trials with T = 10, 15, and 30 minutes) and used the new fitted coefficients to compare predicted vs. observed fractional casualties for the same subset of trials. We also examined measures of goodness of fit. Figure 2 shows an example plot of predicted vs. observed fractional casualties for the subset of 20 trials with T = 10, 15, and 30 minutes, along with a table showing the fitted model coefficients for different subsets of the data and the associated measures of goodness of fit.

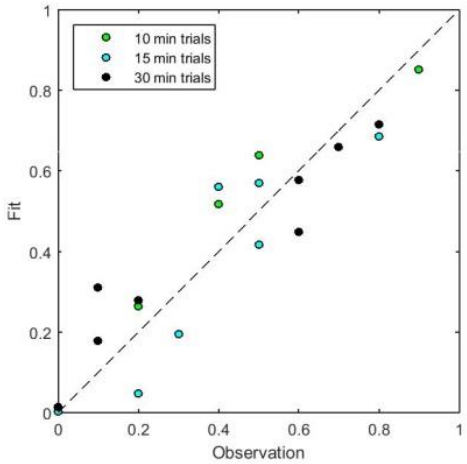

Figure 2. Predicted vs. observed fractional casualties for all the subset of 20 steady-exposure trials with T = 10, 15, and 30 minutes. The predictions used toxic load model coefficients that were fitted using the same subset of data.

Table 1. Toxic load model coefficients and goodness-of-fit metrics (root-mean-square error (RMSE) and *p*-value) derived from fits to different exposure-duration data subsets. The user manual for the US EPA's BMDS software, which was used to make the fits, suggests that $p < 0.1$ indicates a poor fit. Larger RMSE values indicate poorer fits.

Exposure durations (min)	**# of trials**	***n***	TL_{50}	***m***	**RMSE**	***p*-value**
2.3, 5, 10, 15, 30	34	1.35	5.41×10^4	3.87	0.187	0.000337
2,3, 10, 15, 30	27	1.23	2.71×10^4	7.04	0.124	0.221
5, 10, 15, 30	27	1.73	5.20×10^5	3.71	0.145	0.257
10, 15, 30	20	1.36	5.62×10^4	6.15	0.105	0.856
5, 15, 30	20	1.79	7.06×10^5	3.85	0.137	0.351
2.3, 5, 10	19	1.12	1.05×10^4	4.00	0.209	0.005

Figure 2 indicates that the toxic load model is poorly fit when all 34 trials (i.e., all exposure timescales) of the ECBC/NAMRU-D steady-exposure data set are used. The model fit improves, as does the model predictive performance, when data from some exposure durations are removed. In particular, removing the 5-minute exposure time data results in the largest improvement in the fit, and removing the data for the two shortest exposure times (2.3 minutes and 5 minutes) results in the best fit. This suggests that the toxic load model does not perform well across the full set of exposure durations represented in the ECBC/NAMRU-D steady-exposure data set. The toxic load model cannot predict casualties well at both the shorter (2.3 and 5 minutes) and the longer (10, 15, and 30 minutes) exposure durations using the same set of model coefficients. Since HCN exposure durations of both minutes and tens of minutes can result in a significant likelihood of death if the HCN concentration is high enough, this suggests that there may be an intrinsic problem with the toxic load model – at least for HCN inhalation exposures in rats.

In work not reported here, we also calculated *p*-values using an alternative Monte Carlo sampling method since the BMDS-calculated scores from the Pearson's χ^2 test (which we converted to *p*-values in Table 1) can be inaccurate when sample sizes are small (e.g., 10 rats per trial). Although the Monte Carlo-based *p*-values can differ significantly from the BMDS-based ones, our overall findings about which subsets of the data give good fits are not changed appreciably. We further explored the effects of small samples by performing statistical tests to examine whether the scatter in the data could be explained solely by sample size effects. We found that the scatter of the data in the full data set of 34 trials cannot be explained solely by sample size effects, but that the scatter in the T = 10, 15, and 30 minute subset is within the expected scatter. Therefore, we confirm that the 2.3- and 5-minute data are not well-explained by the toxic load model, either due to a systematic experimental error or some physiological process in the rat.

CONCLUSIONS

Our analysis of the ECBC/NAMRU-D data indicates that the toxic load model is not suitable for describing the steady-exposure data over the full range of the experiments' exposure durations (2.3 minutes to 30 minutes): the model fits the data well only if the short-duration exposures (less than 10 minutes) are dropped from the data set. Our analysis does not attempt to attribute a physical explanation for these results. We note that a practical toxicology model should be able to describe toxicological effects across the full timescale of interest (i.e., from minutes to tens of minutes or longer).

Acknowledgements: This effort was supported by the US Defense Threat Reduction Agency (through Mr. Richard J. Fry as the project monitor) and the Institute for Defense Analyses' professional development program. The views expressed in this paper are solely those of the authors.

REFERENCES

EPA, 2015: Benchmark Dose Software (BMDS) User Manual (BMDS 2.6.0), US Environmental Protection Agency, Washington, DC, USA, 339 pp.

Finney, D.J., 1947: Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve. Cambridge University Press, Cambridge, UK, 256 pp.

Platt, N., W.R. Kimball and J.T. Urban, 2014: The use of probabilistic plume predictions for the consequence assessment of atmospheric releases of hazardous materials. *Int. J. of Environment and Pollution*, **55**, 3-12.

Sommerville, D.R., K.H. Park, M.O. Kierzewski, M.D. Dunkel, M.I. Hutton and N.A. Pinto, 2006: "Toxic Load Modeling" in Inhalation Toxicity, 2nd ed., eds. H. Salem and S.A. Katz, CRC Press, New York, NY, USA, pp 137-158.

Sweeney, L.M., D.R. Sommerville and S.R. Channel, 2014: Impact of non-constant concentration exposure on lethality of inhaled hydrogen cyanide. *Toxicol. Sci.*, **138**, 205-216.

Sweeney, L.M., D.R. Sommerville, S.R. Channel, B.C. Sharits, N.M. Gargas and C.P. Gut Jr., 2015: Evaluating the validity and applicable domain of the toxic load model: Impact of concentration vs. time profile on inhalation lethality of hydrogen cyanide. *Regul. Toxicol. Pharmacol.*, **71**, 571-584.

ten Berge, W.F., A. Zwart and L.M. Appelman, 1986: Concentration–time mortality response relationship of irritant and systemically acting vapours and gases. *J. Hazard. Mater.*, **13**, 301-309.

Urban, J.T., M.A. Ambroso, K.J. Galvin, S.M. Nunes, D.F. DeRiggi and N. Platt, 2011: The role of atmospheric dispersion modelling in modern consequence assessment studies. *Proceedings of the 14th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, 713-717.

Urban, J.T., K.J. Galvin, N. Platt, P.E. Bieringer, G. Bieberbach and A.J. Annunzio, 2014: Comparison of hazard area and casualty predictions of a small-scale chemical attack using various toxic load toxicity models. *Int. J. of Environment and Pollution*, **54**, 222-232.

Wilson, D.J., 1995: Concentration Fluctuations and Averaging Time in Vapor Clouds, Center for Chemical Process Safety of the American Institute of Chemical Engineers, New York, NY, USA, 181 pp.