EPA/600/R-03/039
September 1999
Updated October 2000

Research Report:

The Risk Assessment of Mixtures of Disinfection By-products (DBPs) in Drinking Water

by

Linda K. Teuschler (Project Lead), Richard C. Hertzberg, and John C. Lipscomb
National Center for Environmental Assessment
Cincinnati, OH 45268

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Cincinnati, OH 45268

DISCLAIMER

This document has been reviewed in accordance with the U.S. Environmental Protection Agency's peer and administrative review policies and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

FOREWORD

This report contains information concerning research sponsored by the U.S. Environmental Protection Agency's (EPA) National Center for Environmental Assessment - Cincinnati Office (NCEA-Cin) on the risk assessment of mixtures of disinfection by-products (DBPs) in drinking water. Under 42 USC § 300 of the Safe Drinking Water Act Amendments of 1996, it is stated that the Agency will "develop new approaches to the study of complex mixtures, such as mixtures found in drinking water..." This report reflects current results related to research in this area over the past five years. A number of manuscripts and reports covering laboratory research, statistical models, risk assessment methods, and research directions are presented.

NCEA-Cin has been conducting and supporting statistical and biological research pertaining to mixtures of DBPs since 1994 in concert with the Agency's National Health and Environmental Effects Research Laboratory (NHEERL), through several cooperative agreements and in-house. Laboratory scientists at NHEERL performed toxicity experiments in female CD-1 mice. Collaborative statistical research was conducted by Virginia Commonwealth University under cooperative agreements #CR-822671, #CR-820847, and #CR-822517. Laboratory toxicity experiments in Japanese medaka (*Oryzias latipes*) fish embryos were performed by Tulane University under cooperative agreement #CR-822766. Risk characterization methods were developed by NCEA-Cin scientists and by Syracuse Research Corporation under cooperative agreement #CR-822761.

The common goal of all these projects was to determine the systemic toxicity and carcinogenicity of mixtures of DBPs for use in the evaluation of any potential human health risks. The laboratory research was designed for optimal use in assessing critical risk assessment issues, such as characterizing the potency and nature of toxic interactions among DBPs, investigating whether additivity assumptions are useful for DBP risk characterization, providing information on cross-species extrapolations, and estimating health risks associated with different drinking water treatment options. Thus, the multiple-purpose design approach of this research combined efficient laboratory experimental designs with statistical models to provide data on critical research issues and validate new risk characterization methodologies.

RESEARCHERS

This research on the risk assessment of disinfection by-products was jointly sponsored by NCEA-Cin's Comparative Risk Project Team and Mixtures Research Team. The scientists who conducted this research and are the primary authors of the material in this report were as follows (listed alphabetically):

National Center for Environmental Assessment - Cincinnati Division, Cincinnati, OH
Richard C. Hertzberg
John C. Lipscomb
Linda K. Teuschler (Project Lead)

National Health and Environmental Effects Research Laboratory,
Research Triangle Park, NC
Tony McDonald
Yusupha Sey
Jane Ellen Simmons

Syracuse Research Corporation, Syracuse NY
Joan Colman
William M. Stiteler

Tulane University Medical Center, New Orleans, LA
William Hartley
Rani Thiyagarajah

Virginia Commonwealth University, Richmond, VA
Hans Carter
Chris Gennings
Pam Schwartz (now with Pfizer, Inc., Groton, CT)

EXTERNAL REVIEW OF RESEARCH

This research program was reviewed by an expert panel in May 1998, with a workshop report generated (U.S. EPA, 1998). The research presented in this report was greatly enhanced by responses and questions raised by the workshop panel. The expert panel who critiqued the research program and recommended future directions included the following members:

Gary A. Boorman
NIEHS
Laboratory of Experimental Pathology
National Toxicology Program
Research Triangle Park, NC

Tracie E. Bunton
Johns Hopkins University
Division of Comparative Medicine
Baltimore, MD

James J. Chen
National Center for Toxicological Research
Food and Drug Administration
Jefferson, AR

Christopher Cox
University of Rochester
Division of Biostatistics and Division of Toxicology
School of Medicine and Dentistry
Rochester, NY

Vicki Dellarco
U.S. Environmental Protection Agency
Office Of Water
Washington, DC

Victor Feron
TNO Nutrition and Food Research
Utrechtseweg
The Netherlands

Kannan Krishnan
University of Montreal
Dept. Of Medicine
Montreal, Canada

Mike Pereira
Medical College of Ohio
Department of Pathology
Toledo, OH

TABLE OF CONTENTS

TABLE OF CONTENTS (cont.)

INTRODUCTION

The U.S. Environmental Protection Agency's (EPA) National Center for Environmental Assessment-Cincinnati Division (NCEA-Cin) has conducted and supported statistical and biological research pertaining to mixtures of disinfection by-products (DBPs) since 1994 in-house, in concert with the Agency's National Health and Environmental Effects Research Laboratory (NHEERL), and through several cooperative agreements. Research has been conducted in collaboration with Virginia Commonwealth University (VCU) and NHEERL to generate efficient experimental designs to test specific mixtures for departures from additivity by building joint, dose-response threshold additivity models using single chemical data. To test the usefulness of these statistical models, NHEERL laboratory scientists generated health effects data for hepatotoxicity and nephrotoxicity in female CD-1 mice exposed to mixtures of DBPs, specifically trihalomethanes (THMs). Mixtures toxicity data were also developed for the THMs by Tulane University (Tulane) on carcinogenic, developmental, and reproductive effects in Japanese medaka (*Oryzias latipes*) fish embryos. In addition, NCEA-Cin and Syracuse Research Corporation (SRC) investigated the development of risk assessment methodologies to characterize the health risks from exposure to mixtures of DBPs across various treatment scenarios. The laboratory experiments in the mice and in the medaka were designed to provide data useful for a number of purposes, including the testing and refinement of statistical and risk assessment methods, as well as develop an understanding of the toxicity and potential interactions of mixtures of DBPs. The specific goals of this research were the following:

- To investigate priority chemicals
- To establish or refute additivity at low doses
- To use threshold models for multiple chemicals
- To develop efficient experimental designs for mixtures
- To use binary and single chemical data to estimate mixtures toxicity
- To investigate relevant endpoints
- To corroborate effects across species
- To evaluate the medaka as a screening bioassay for mixtures
- To compare typical mixing ratios for an ozonation process and a chlorination process
- To estimate health risks for multiple DBPs for typical treatment processes

This research has been presented at professional meetings, published, and reviewed on many occasions and in various formats (see Appendix A for a complete list of citations). On May 19, 1998, this group of researchers presented their data and methods to an expert panel (see page v of this report) during a scientific review workshop in Cincinnati, Ohio. The panel's report (U.S. EPA, 1998) lauded the efforts of this research, stating, "The panel would like to applaud EPA for the collaborative approach that has been taken with NCEA, NHEERL, Tulane, Syracuse Research Corporation, and Virginia Commonwealth University..... We encourage sharing of data, tissues, etc., across these different programs." The workshop panel made many important observations and recommendations that greatly influenced and enhanced subsequent research. The material presented in this report represents the progress made since the 1998 workshop in new data collection, statistical analysis, and methods development.

This report reflects the current results related to research in this area over the past 5 years by presenting a number of manuscripts and reports covering laboratory research, statistical models, risk assessment methods, and research directions. The need to study DBP mixtures is mandated under 42 USC § 300 of the Safe Drinking Water Act Amendments of 1996, where it is stated that the EPA will "develop new approaches to the study of complex mixtures, such as mixtures found in drinking water..." In addition, the EPA's Office of Water drafted a *Research Plan for Microbial Pathogens and DBPs in Drinking Water* that calls for the characterization of DBP mixtures risk (http://www.epa.gov/water). Trihalomethanes in particular are important to study because they are found in relatively large concentrations and are regulated as a mixture (limited to a total of 80 µg/L in finished drinking water). The results of this research are valuable to drinking water health risk evaluation and to the broader field of mixtures research. Many aspects of this work, such as the development of efficient experimental designs and laboratory methods for mixtures and the generation and refinement of new risk characterization approaches, are transferable to the evaluation of other environmental mixtures.

These papers and reports are arranged so those dealing with the general approach, rationale, and concepts are presented first; those that lay out laboratory methods and toxicity data results are next; and those that present new statistical or risk assessment approaches to data analysis are last. It may be noted that some of the early efforts to develop new statistical and laboratory methods involved chemicals other than DBPs (i.e., benzene studies in the medaka presented in Section 6 and a study using tetrahydroaminoacridine to illustrate the initial development of the threshold model shown in Section 12. These papers are included as contributors to the subsequent DBP evaluations. Several of the papers in this document are in various stages of publication (see Appendix A for citations). Two of these papers on the toxicologic and statistical aspects of this work (Sections 4 and 5, respectively) have been published as companion papers in the *1999 American Water Works Association Conference Proceedings*. The overview paper in Section 3 has been published in *Drug and Chemical Toxicology*. The paper on experimental designs for estimation of thresholds (Section 12) has been accepted for publication by the *Journal of Agricultural, Biological and Environmental Statistics*. The paper on proportional response addition has been submitted to *Risk Analysis* (Section 11). Each section is written as a stand-alone document, but taken as a whole, they present a comprehensive overview of this research program.

Appendix B has been added to this document to correct an error discovered in the dosing values of chlorodibromomethane (CDBM) used in the mixture analyzed by a threshold additivity model in Gennings et al. (1997). This paper is cited throughout this report, and the modeling results have been referenced in both the paper on novel statistical methods (Section 5) and in the description of the multiple design approach (Section 3). Appendix B presents the errata to the original Gennings et al. (1997) paper. The errata was published in 2000 in the *Journal of Agricultural, Biological and Environmental Statistics*. The errata provides corrected tables and figures for that manuscript as well as their interpretation. The doses used in the single chemical experiments for each of the four trihalomethanes (THMs) tested (of which CDBM is one) are given as 0, 0.152, 0.305, 0.76, 1.52, and 3.05 mmol/kg/day. For CDBM, in the single chemical studies only, the correct doses are 0, 0.304, 0.610, 1.52, 3.04, and 6.10 mmol/kg/day. Although the errata updates the additivity modeling results, these new data continue to support the bottom

line conclusion of this study of no evidence of departure from additivity at the combination point of interest. The progress report on interactions among the trihalomethanes (Section 8) does reflect these corrected data. Summaries of the papers and progress reports found in this document follow.

Section 2: Approaching the Toxicity of Disinfection By-products in Drinking Water as a Mixtures Problem
(L.K. Teuschler, J.E. Simmons)

Human health risk assessment from exposure to DBPs in drinking water is of concern because of the wide spread exposure to persons who receive disinfected water. Taken as a body of literature, epidemiologic studies on chlorinated drinking water offer some evidence of an association with certain cancers and reproductive and developmental effects, such that further investigations are warranted. In addition, evidence exists of mutagenicity in *in vitro* studies of drinking water extracts and evidence of carcinogenicity, reproductive effects, nephrotoxicity and hepatotoxicity in *in vivo* single chemical studies at high doses of DBPs. Because typical human exposures to DBPs are to extremely low levels of these chemicals, it is hypothesized that any health effects risk is likely to be attributable to factors other than exposure to any single DBP. Three approaches are suggested for the evaluation of DBP mixtures, based on expert panel recommendations in an International Risk Science Institute report (ILSI, 1998). These include toxicologic studies of simple defined mixtures, toxicologic studies using reproducible disinfection scenario samples, and toxicologic or epidemiologic studies on direct drinking water samples. This paper suggests that all three approaches are valid and can provide data applicable to the risk assessment of DBPs. Furthermore, it is posited that high-dose, single chemical animal toxicity studies alone are not adequate for characterizing DBP health risks. A concentrated effort by a multidisciplinary team of researchers is necessary to generate the appropriate data for this problem.

Section 3: A Multiple-Purpose Design Approach to the Evaluation of Risks from Mixtures of Disinfection By-Products
(L.K. Teuschler, C. Gennings, W.M. Stiteler, R.C. Hertzberg, J.T. Colman, A. Thiyagarajah, J.C. Lipscomb, W.R. Hartley, J.E. Simmons)

Researchers from several organizations have developed a multiple-purpose design approach to the evaluation of DBPs mixtures that combines efficient laboratory experimental designs with statistical models to provide data on critical research issues (e.g., estimation of human health risk from low-level DBP exposures, evaluation of additivity assumptions as useful for risk characterization, estimation of health risks from different drinking water treatment options). A series of THM experiments have been designed to study embryonic development, mortality, and cancer in Japanese medaka (*Oryzias latipes*) and liver and kidney endpoints in female CD-1 mice. The studies are to provide dose-response data for specific mixtures of the four THMs, for the single chemicals, and for binary combinations. The dose-levels and mixing ratios for these experiments were selected as useful for development and refinement of three different statistical methods: testing for departures from dose-additivity; development of an interactions-based hazard index; and use of proportional-response addition as a risk

characterization method. Preliminary results suggest that dose-additivity is a reasonable risk assessment assumption for DBPs. The future of mixtures research depends on such collaborative efforts that maximize the use of resources and focus on issues of high relevance to the risk assessment of human health.

Section 4: Advances in the Toxicological Assessment of Disinfection By-Products in Rodent and Fish Biomedical Models
(W.R. Hartley, C. Gennings, L.K. Teuschler, A. Thiyagarajah, J.E. Simmons)

Determination of the toxicological interactions of DBP mixtures is essential in the risk assessment process and selection of drinking water disinfection methods which minimize potential health risks. Efficient and short-term toxicological methods for assessing the toxicity of DBP mixtures are presented, including the use of the Japanese medaka (*Oryzias latipes*) for assessment of developmental toxicity (neurological, circulatory/heart, and reproductive effects) and cancer and the rodent for assessment of hepatic and renal toxicity. The proposed methods for the medaka and rodent are demonstrated by toxicological evaluation of bromoform and chloroform. As these toxicological methods are further refined, they may be used to provide toxicological information for risk management decisions to reduce health risks from drinking water with DBP contaminants.

Section 5: Novel Statistical Methods for Risk Assessment of Disinfection By-Product Mixtures
(C. Gennings, L.K. Teuschler, W.R. Hartley, A. Thiyagarajah, J.E. Simmons)

Investigation of the assumption of additivity has been hampered by the lack of adequate and appropriate experimental designs and statistical methods. This is particularly true for mixtures of more than two chemicals. Here, we describe the threshold additivity model (Gennings et al., 1997), a flexible experimental design and statistical methodology applicable to mixtures of large numbers of chemicals. Advantages of this design and analytic approach are that it allows investigators to focus on particular mixture combination points of interest, decreasing the number of animals required as well as the time and cost of performing the experiments. This approach is illustrated by two examples applicable to drinking water disinfection by-products formed during either chlorination or ozonation of water followed by post-treatment with either chlorine or chloramine. In the first example, the additivity assumption is examined, with hepatotoxicity in female CD-1 mice as the endpoint, for a mixture of the four trihalomethanes, chloroform ($CHCl_3$), bromoform ($CHBr_3$), bromodichloromethane, and chlorodibromomethane, all formed during disinfection of water. The mixture tested was based on the average seasonal proportions of these four chemicals at 35 water treatment facilities (Krasner et al., 1989). For the particular mixture and dosage tested, the experimental sample mean was within the 95% prediction interval from the threshold additivity model, providing evidence that dose additivity is a reasonable assumption for risk assessment. In the second example, the additivity assumption for developmental toxicity is examined for binary mixtures of $CHCl_3$ and $CHBr_3$ in medaka fish. With the statistical power afforded by the present experiment, antagonism was detected at the highest mixture dose tested (25 ppm $CHCl_3$: 25 ppm $CHBr_3$) and departure from additivity was not detected at the three lower-dose mixture groups.

Section 6: Progress Report: Carcinogenicity of Benzene in the Japanese Medaka (*Oryzias Latipes*) Exposed During Embryological Development
(W.R. Hartley, A. Thiyagarajah)

Benzene is a volatile organic compound, found in contaminated air, water, sediment, soils, food, and tobacco smoke. Benzene is classified as a known human carcinogen by U.S. EPA and has produced an array of tumor types in rodent studies. The goals of this study were to determine the carcinogenic effects of benzene in Japanese medaka using a 10-day embryonic exposure methodology, to contrast the results from the standpoint of cross-species comparisons, and to discuss potential use of these methods for risk assessment. In this exploratory research, the basis for limiting exposure periods to the embryo were known sensitivity of the medaka to the carcinogens in long-term exposure; transparency of the chorionic membrane (membrane surrounding the developing embryos) so developmental effects could be continuously or periodically observed; simplicity of exposure regimens; and the ability to focus on critical periods of development. Medaka embryos were exposed to 0, 43.5, 100, and 150 mg/L benzene beginning at early high blastula for ten continuous days.

Results showed that the only liver tumors noted in one-year-old medaka exposed as embryos to benzene were hepatic adenoma and cholangioma. There was an increasing trend ($p \leq 0.1$) in the total tumor incidence in fish surviving after one year in the two high concentration (100 and 150 mg/L benzene) groups, with one case of hepatic cholangioma in the control group. Based upon the known sensitivity of the medaka to known carcinogens in lifetime (chronic) exposure studies and frequent use in cancer research (Hawkins et al., 1985; Chen et al., 1996), this study used an embryo exposure route to determine the sensitivity of the medaka to the carcinogenicity of benzene. Because the exposure period was limited to 10 days of embryogenesis and new methods were being used, high (shock) concentrations of benzene of known embryo lethality was administered to optimize penetration of benzene through the chorion. The hepatocellular adenomas and cholangiomas observed in this study were benign neoplasms. The malignant neoplasms associated with benzene exposure in the previously summarized rodent literature were not observed in this study. Considering the marginal statistical significance of the increased liver tumors and the absence of malignant liver tumors, this medaka embryo exposure methodology does not indicate the known carcinogenic potential of benzene and does not reflect the highly variable carcinogenic response reported in numerous rodent bioassays. The low incidence of non-neoplastic proliferative and degenerative lesions observed in this study are probably a consequence of normal aging. To further evaluate the utility of the embryo exposure methodology for risk assessment, a similar protocol will be tested for the potent animal carcinogen diethylnitrosamine (DEN). The exposure period may also be extended to include young fish in the range of several weeks after hatching.

Section 7: Progress Report: Chronic Toxicity of the Disinfection By-Products (DBP) Chloroform and Bromoform in the Japanese Medaka (*Oryzias Latipes*)
(W.R. Hartley, A. Thiyagarajah)

The trihalomethanes (THMs), chloroform, bromoform, bromodichloromethane (BDCM), and chlorodibromomethane (CDBM), are present in drinking water in relatively large quantities,

compared with other DBPs. Because epidemiologic studies suggest the potential for cancer and developmental effects in humans from DBP exposures, THMs were selected for evaluation. Japanese medaka embryos were exposed for the first 10 days of embryological development to chloroform (25, 50, or 100 ppm) or bromoform (5, 10, 25, or 50 ppm). Two control groups were used, one in an embryo rearing solution (ERS) and one in ERS/dimethyl sulfoxide (DMSO). At 6 and 12 months, the fish were evaluated for chronic effects including cancer. Preliminary results are given in this progress report.

For chloroform, a low incidence of neoplasms in the liver was observed. After 12 months, cholangioma was found in two fish at 25 mg/L and one fish at 100 mg/L, and lipoma was found in one fish at 50 mg/L. Non-neoplastic proliferative lesions observed were clear cell foci, vacuolated cell foci, and altered hepatocellular altered foci. Excessive accumulation of substances such as glycogen or lipid occurred in large numbers of fish after 6 months; the incidence decreased in fish examined after 12 months. In addition, hepatic cysts and spongiosis hepatis were found in high incidence. In the kidney, the incidence of cystic tubules in the control and exposed fish were similar. Other tubular lesions observed were accumulation of hyaline globules in the tubular epithelium and a lower incidence of vacuoles in the tubular epithelium and tubular necrosis. In the thyroid, infollicular hyperplasia increased in the fish exposed for 12 months. Delayed sexual maturation was observed at all concentrations including controls. Immature testis were found only in fish examined after 6 months; none of the 12-month-old fish had immature testis except one control fish. The incidence of immature testis in fish examined after 6 months was higher than controls. In female fish, immature ovaries occurred in all groups including controls, with the incidence higher in the exposed fish.

For bromoform, the incidence of liver lesions was similar to that observed in chloroform-exposed fish. Incidence of hepatic cysts was high compared to controls. One fish exposed at 50 mg/L for 6 months had cholangioma. In the kidney, a high incidence of cystic tubules occurred in all groups of fish including controls. In the thyroid, a high incidence of follicular hyperplasia occurred in 12-month-old fish including controls as compared to 6-month-old fish. A high incidence of immature testis and ovary occurred in 6-month-old fish exposed at 10 mg/L.

At this time (September 1999), this progress report provides the largely complete laboratory results on chloroform and significant information on bromoform. Statistical analyses have not yet been performed on these data. Completion of the 12-month data sets for bromoform at 10 mg/L and DMSO are in progress. The 50 mg/L bromoform group was terminated prior to 12 months because of high mortality. The 5 and 25 mg/L bromoform groups (6 and 12 months) are currently under histological evaluation. Similar chronic experiments for BDCM are in progress. This progress report does not include an update on cardiac lesions (dilated atrium and tubular heart) as special sectioning of the tissues is being performed to more clearly understand the etiology of the high incidence of these lesions.

Section 8: Progress Report: the Hepatotoxic Interaction of Chloroform, Bromodichloromethane, Bromoform and Chlorodibromomethane
(J.E. Simmons, C. Gennings, A. McDonald, P.F. Schwartz, W.H. Carter, Jr., L.K. Teuschler, Y.M. Sey)

Assessment of the nature (additive, synergistic, or antagonistic) of the interaction between the trihalomethanes (THMs), chloroform ($CHC1_3$), bromodichloromethane (BDCM), chlorodibromomethane (CDBM) and bromoform ($CHBr_3$), provides useful information in assessing the human health risks associated with disinfection of water. A threshold additivity model was developed for this purpose (Gennings et al., 1997); in this method, the response at specific mixture concentrations of interest is predicted under dose addition by use of data from the single chemical dose-response curves for each chemical in the mixture. The concentrations of the four THMs mixtures tested at total doses of 0.436 and 0.872 mmol/kg/day were based on the average seasonal proportions at 35 water treatment facilities (Krasner et al., 1989). Female CD-1 mice, ~65-70 days old, were gavaged daily for 14 days with this mixture in 10% Alkamuls®. Based on the single chemical dose-response curves generated in our lab, both the serum sorbitol dehydrogenase (SDH) and alanine aminotransferase (ALT) levels predicted by the threshold additivity model for both mixtures fell within 95% prediction intervals of the response if dose-addition is a good assumption. The closeness of the predicted mixture responses and the observed mixture response indicates that the four chemicals are dose-additive at the mixture combination tested.

Section 9: Progress Report: the Effect of Mixing Ratio on Mixtures of Trihalomethane Disinfection By-Products
(A. McDonald, C. Gennings, W.R. Hartley, L.K Teuschler, A. Thiyagarajah, Y.M. Sey, S.W. Krasner, J.E. Simmons)

Water treatment processes that involve ozonation/chloramination decrease the total amount of THMs formed but alter their relative distribution. The objective of this investigation was to determine the effect of mixing ratio on mixtures of the four THMs. Female CD-1 mice (~65-70 days old) were gavaged daily for 14 days with the THM mixtures in 10% Alkamus® El-620. Hepatotoxicity was assessed on day 15 by serum sorbitol dehydrogenase (SDH) and by histopathology. Relevant mixing ratios were selected from a study in which high bromide water was treated either by prechlorination/postchlorination (CL) or by preozonation/post-chloramination (OZCM). The proportions of the CL mixture were $CHCl_3$, 0.319; $CHBr_3$, 0.049; BDCM, 0.342; and CDBM, 0.290. The proportions of the OZCM mixture were $CHCl_3$, 0.655; $CHBr_3$, 0.029; BDCM, 0.187; and CDBM, 0.129. Three experiments were conducted, each testing a different total dosage of these two mixtures: 0.05, 1.50, and 3.0 mmol/kg/day. At each total mixture dosage, the SDH values for the CL and OZCM mixtures did not differ significantly. Similarly, the proportion of mice with centrilobular hepatocellular necrosis did not differ significantly between the two mixing ratios. In conclusion, at the same dosage, mixing ratio, or at least the two mixing ratios tested here, had little or no apparent effect on the hepatotoxicity of THM mixtures. Data such as these may help in selection of optimal water treatment strategies. As the total mixture dose was held constant in each experiment, differences

resulting from the amount of THMs produced by each treatment are not reflected but are worthy of future investigation.

Section 10: Categorical Regression Analysis of Bromodichloromethane Liver Toxicity and Pathology Data
(L.K. Teuschler, W.R. Hartley, A. Thiyagarajah, J.C. Lipscomb, J.E. Simmons)

A categorical regression (CR) procedure is being investigated to compare new toxicity data with existing benchmarks, such as a Reference Dose (RfD), and to predict pathology scores from other hepatotoxic endpoints. CR was used to model data on liver toxicity and histopathology scores from female CD-1 mice exposed for 14 days by gavage to doses of 0, 0.152, 0.305, 0.76, 1.52, and 3.05 mmol/kg/day bromodichloromethane (BDCM), a common by-product of drinking water disinfection processes. The toxic effects included both continuous endpoints (liver enzyme elevations, liver weight changes) and categorical data (centrilobular necrosis severity designations). The centrilobular necrosis data were classified into ordered categories of severity (i.e., category 1 = none detected, category 2 = mild, mild-moderate to moderate, and category 3 = moderate-severe to severe) at the individual animal level. Two different CR models were then constructed. The *first CR model* was used to predict the probability of centrilobular necrosis at the Lowest-Observed-Adverse-Effect Level (LOAEL) used to estimate the BDCM RfD. At the animal chronic LOAEL of 17.9 mg/kg/day based on renal cytomegaly in male mice, used to calculate the RfD, the probability of observing mild centrilobular necrosis or greater was 0.01 (95% CI of 6E-4, 0.14), which is consistent with the fact that the kidney effect should be a more sensitive endpoint. The *second CR model* was applied to a test data set of eight mice dosed with 1.2 mmol/kg/d BDCM to predict the occurrence and severity of centrilobular necrosis based on information on the liver serum endpoint, alanine aminotransferase (ALT) level, and the relative liver weight. These predictions were then compared with the actual pathology scores. For occurrence, seven of the eight observed pathology severity scores matched the prediction of at least mild centrilobular necrosis. For severity, in four of the eight mice, the observed pathology severity score corresponded exactly to the most likely severity as predicted by the model. The CR procedures demonstrated here are a useful way to evaluate categorical and qualitative histopathology data.

Section 11: A Proportional-Response Addition Model for Evaluating Health Risks Associated with Complex Mixtures: an Example Using Drinking Water Disinfection By-Products
(W.M. Stiteler, L.K. Teuschler, J.T. Colman, R.C. Hertzberg)

Complex mixtures, such as those resulting from disinfection of drinking water supplies, present a difficult problem for risk assessors because of the large number of components. The concentration of each individual component by itself may be small enough to pose no threat, but if they act jointly in some additive or greater than additive manner, then health safety may be a concern. For the most part, models for the joint action of chemical mixtures are based on assumptions about the mechanisms of action. For example, response addition is based on the assumption of independent but joint action, while dose addition is based on an assumption of simple similar action. This reliance on mechanism becomes problematic for complex mixtures.

As the number of components becomes large, it is unlikely a single mechanism-based definition of additivity would be universally applicable across the whole mixture. There are, for example, 1225 different pairs of components in a 50-component mixture. Different pairs of components may exhibit different modes of action. Even if a common mode of action can be assumed across the mixture, the usual definitions of additivity may not be appropriate. If all components are present at a subthreshold level, then response addition will indicate no response for the mixture, regardless of the number of components. The solution proposed in this paper is to use proportional-response addition, a generic definition of additivity that does not depend on mechanism of action, for modeling complex mixtures. An example of the application of this model to a mixture of DBPs is presented.

Section 12: Optimizing the Precision of Toxicity Threshold Estimation Using a Two-Stage Experimental Design
(P.F. Schwartz, C. Gennings, L.K. Teuschler, M.W. Fariss)

An important consideration for risk assessment is the existence of a threshold: the highest toxicant dose where the response is not distinguishable from background. A methodology has been developed for finding an experimental design that optimizes the precision of threshold model parameter estimation for single-chemical threshold experiments. Being interested in precisely estimating the threshold parameter, D-optimality and D_s-optimality criteria for parameter estimation have been used. The D-optimal design results in parameter estimates as precise as possible in the sense that the likelihood-based confidence region has minimum volume, while the D_s-optimal threshold design results in parameter estimates as precise as possible in that the variance of the threshold parameter estimate is minimized. For nonlinear models, optimal designs are a function of the unknown parameters via the information matrix. Therefore, estimates of the parameters must be obtained before the optimal design of the experiment can be found. For this reason, a two-stage, D-D_s optimal design is recommended where the D-optimality criterion is used in the first stage followed by the D_s-optimality criterion in the second stage. The first stage is used for range finding and to obtain good global estimates to supply to the second stage. The second stage results in precise parameter estimates with minimum variance for the threshold parameter estimate. We propose that this two-stage D-D_s optimal design will provide toxicologists with the experimental parameters necessary to accurately estimate thresholds for risk assessment purposes in a more cost effective and timely manner.

As discussed above, this collaborative research has been conducted to address multiple statistical, toxicologic, and risk assessment issues with relevance to both DBP mixtures and other multiple chemical exposures. The THMs have been investigated as priority drinking water chemicals whose toxicity and joint action is important to understanding potential human health risks. In CD-1 mice experiments, real world mixing ratios of THMs were tested, comparing an ozonation process and a chlorination process and showing no difference in the toxicity. However, further testing is needed using chemicals other than the THMs, such as bromate, that are uniquely or more typically produced by one treatment process or the other. Continued development of these data will aid in the estimation of health risks from exposure to multiple DBPs for specific treatment processes. Evidence of additivity at relatively low doses has been

produced by applying the threshold additivity model to laboratory data in CD-1 mice and in the medaka. Such evidence is important to risk assessment because it supports the use of assumptions of additivity as default procedures in the risk assessment of multiple DBPs at low doses. Because hundreds of DBPs exist that are potential candidates for toxicologic testing, such efficient experimental designs are necessary, and inexpensive bioassays for the screening of combinations of chemicals are essential. In these studies, use of the medaka fish has been explored for screening effects in mixture studies, such as cancer, developmental defects, and systemic effects (e.g., liver, kidney, thyroid, and cardiac effects). Results for the medaka are preliminary, as the protocols for the testing procedures are still being developed; however, use of the medaka as a screening tool is promising. This is particularly true because cancer and developmental effects, the endpoints of most concern in the drinking water epidemiologic literature, are relevant endpoints observed in the medaka experiments. Collaboration on the design of the CD-1 mice experiments and the medaka was done so investigators can corroborate effects across species.

Of particular interest is the comparison of liver enzyme levels and hepatotoxic histopathology results for comparable dose levels between the two species, which can be used to evaluate the medaka bioassay for use in screening mixtures for further study in *in vivo* experiments. Finally, the binary experiments underway in the CD-1 mice have been designed not only for use in the development of threshold additivity models, but also to test the proportional response addition model (Sections 3 and 11) and the interaction-based hazard index approach (Section 3). Both of these methods require experimental data to test their relevance for use in human health risk assessment and to refine the models and parameters being proposed. Some of this testing has begun using the binary information and whole mixture data currently available, with results anticipated when all the laboratory studies are completed.

DBP mixtures research will continue over the next few years to generate additional data and further refine the statistical and risk assessment methods being developed. Future directions potentially include the generation of laboratory data on defined mixtures of not only the THMs, but also the haloacetic acids (HAAs), mixtures of THMs and HAAs together, the haloacetonitriles (HANs), and other DBPs of concern, such as bromate. Statistical methodology is being developed allowing simultaneous inference regarding additivity for multiple mixtures of interest with a specified Type I error rate. In addition, carefully designed experiments should be undertaken to estimate the toxicity of complex mixtures of DBPs produced by using reproducible disinfection scenarios with controls for treatment train and source water characteristics. Using these data in conjunction with the data on defined mixtures, estimates can then be made of the potential toxicity of unidentified DBP material in finished drinking water.

Another area of importance is the determination of the toxic mode(s)-of-action of the DBPs and the mechanism(s) underlying any identified nonadditive toxicity and their relevance to humans for cancer, developmental, and reproductive effects. Mode-of-action is difficult to determine, but is critical not only to understanding toxicity, but also to the development of statistical risk models based on biologic assumptions (e.g., independence of action, similar toxicologic action). Finally, the development of appropriate laboratory test systems and efficient experimental designs are crucial to this area of research because of the myriad DBPs that have

been identified in finished drinking water. Investigations into short-term screening tests for DBP mixtures, such as *in vivo* systems (e.g., developmental effects/cancer in medaka fish) or *in vitro* procedures (e.g., carcinogenicity in Syrian Hamster Embryo Cell Transformation Assay; developmental effects in whole embryo culture systems) are needed to identify the potential toxicity of DBP mixtures. The most toxic mixtures can then be tested in appropriate rodent test systems for developmental, reproductive, hepatotoxic, nephrotoxic, carcinogenic, or other effects.

References

Chen, H.C., I.J. Pan, W.J. Tu, W.H. Lin, C.C. Hong and M.R. Brittelli. 1996. Neoplastic response in Japanese medaka and Channel catfish exposed to N-methyl-N'Nitro-N-Nitrosoguanidine. *Toxicol. Pathol.* 24:696-706.

Gennings, C., P. Schwartz, W.H. Carter, Jr. and J.E. Simmons. 1997. Detection of Departures from Additivity in Mixtures of Many Chemicals with a Threshold Model. *Journal of Agricultural, Biological and Environmental Statistics* 2:198-211. (Erratum: *Journal of Agricultural, Biological and Environmental Statistics*. 5(2), 2000).

Hawkins, WE., R.M. Overstreet and W.W. Walker. 1988. Carcinogenicity tests with small fish species. *Aquatic Toxicol.* 11:113-128.

ILSI (International Life Sciences Institute). 1998. Assessing the Toxicity of Exposure to Mixtures of Disinfection By-Products. Research Recommendations. Report prepared by the ILSI Risk, Science Institute under a cooperative agreement with U. S. EPA's Office of Water. Washington, DC.

Krasner, SW., M.J. McGuire, J.G. Jacangelo, N.L. Patania, K.M. Reagan and E.M. Aieta. 1989. The occurrence of disinfection by-products in U.S. drinking water. *JAWWA*. 81:41-53.

U.S. EPA. 1998. Workshop Report: The Risk Assessment of Mixtures of Disinfection By-Products (DBPs) in Drinking Water. NCEA-C-0459.

APPROACHING THE TOXICITY OF DISINFECTION BY-PRODUCTS IN DRINKING WATER AS A MIXTURES PROBLEM

Linda K. Teuschler
National Center for Environmental Assessment - Cincinnati
U.S. Environmental Protection Agency

Jane Ellen Simmons
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency

August, 1999

Abstract

Human health risk assessment from exposure to disinfection by-products (DBPs) in drinking water is of concern because of the wide spread exposure to persons who receive disinfected water. Taken as a body of literature, epidemiologic studies on chlorinated drinking water offer some evidence of an association with certain cancers and reproductive and developmental effects such that further investigations are warranted. In addition, there exists evidence of mutagenicity in *in vitro* studies of drinking water extracts and evidence of carcinogenicity, reproductive effects, nephrotoxicity, and hepatotoxicity in *in vivo* single chemical studies at high doses of DBPs. Because typical human exposures to DBPs are to extremely low levels of these chemicals, it is hypothesized that any health effects risk is likely to be attributable to factors other than exposure to any single DBP. Three approaches are suggested for the evaluation of DBP mixtures, based on expert panel recommendations in an International Life Sciences Institute report (ILSI, 1998). These include toxicologic studies of simple defined mixtures, toxicologic studies using reproducible disinfection scenario samples, and toxicologic or epidemiologic studies on direct drinking water samples. This paper suggests that all three approaches are valid and can provide data that are applicable to the risk assessment of DBPs. Furthermore it is posited that high-dose, single chemical animal toxicity studies alone are not adequate for characterizing DBP health risks. A concentrated effort by a multidisciplinary team of researchers is necessary to generate the appropriate data for this problem.

Introduction

Human health risk assessment from exposure to disinfection by-products (DBPs) in drinking water is of concern because of the widespread daily exposure to this complex mixture. Although it is clear that water disinfection is effective in preventing waterborne microbial illnesses, as early as 1974, the potential health risks from exposure to chemical by-products of disinfection processes were recognized. DBPs are produced when disinfectants such as chlorine, ozone, chloramine, or chlorine dioxide react with naturally occurring organic matter in the water. The most common DBPs on which concentration data are available include the trihalomethanes (THMs), haloacetic acids (HAAs), haloacetonitriles, haloketones, aldehydes, bromate, chloral hydrate, and chloropicrin, among others (Jacangelo et al., 1989; Krasner et al., 1989; Lykins et al., 1994; Miltner et al., 1990). More recently, Richardson (1998) identified several hundred DBPs from various disinfection scenarios. Of the identified DBPs, fewer than 20 have been subjected to toxicity studies of use in risk assessment.

Data from both epidemiologic and toxicologic studies offer some evidence that human health effects from DBP exposure are of concern. Epidemiologic studies of chlorinated drinking water exposures in humans suggest weak associations primarily with bladder, rectal, and colon cancer (Cantor et al., 1985, 1997; McGeehin et al., 1993; King and Marrett, 1996; Freedman et al., 1997) and limited evidence of reproductive and developmental effects (Bove et al., 1995; Kramer et al., 1992; Swan et al., 1998; Waller et al., 1998). In contrast, general toxic effects have not been observed when animals were exposed to finished drinking water in which DBPs occur at extremely low levels (Bull et al., 1982; Kavlock et al., 1979). At high DBP dose levels of single DBPs however, there is toxicologic evidence of carcinogenicity, reproductive effects, developmental effects, and other toxic effects, particularly in the kidney and liver (Bull and Kopfler, 1991; NTP, 1985, 1986, 1989; Smith et al., 1989). There is also evidence of mutagenicity from exposure to extracts of finished drinking water in *in vitro* studies (Kool et al., 1981; Loper et al., 1978; Nestmann et al., 1982). Although there are few studies available on defined mixtures of DBPs, evidence exists of dose additivity for liver effects in mice exposed to mixtures of trihalomethanes (THMs) (Gennings et al., 1997) and of promotion of cancer by mixtures of dichloroacetic acid (DCA) and trichloroacetic acid (TCA) (Pereira et al., 1997).

Risk assessment questions surround the issues of establishing, explaining, and estimating any substantive human health risks from exposure to the low levels of DBPs found in drinking water. Because toxic effects are not observed in animal studies when the exposures are to low doses and because the epidemiologic data are inconsistent across studies with only relatively weak associations noted, the existence of human health risks is questionable, but cannot be entirely dismissed. If it is assumed, however, that the human health effects suggested in epidemiologic studies are real, then several hypotheses can be posed to explain the discrepancies between the epidemiologic results and the lack of effects in animals exposed to finished drinking water. Such hypotheses include the following:

1) there is an effect from exposure to the mixture of DBPs that is at least additive (if not synergistic) in nature, so studies involving low levels of individual DBPs are inadequate to explain the health effects found in the positive epidemiologic data;

2) effects in humans are the result of the chronic, repetitive insult from daily exposure to DBP mixtures;

3) animal studies differ from human exposures in ways such as differences in physiology, biochemistry, anatomy, and lifestyle factors (e.g., high fat diets) that prevent them from demonstrating the same outcomes;

4) laboratory studies to date expose the animals to only a single route, usually oral, so that effects resulting from dermal or inhalation exposure are not observed;

5) effects in epidemiologic studies result from exposure to other environmental factors such as metals and inorganic materials in the drinking water or concurrent exposures to industrial pollutants in urban areas or pesticides in agricultural areas so animal studies solely focused on DBPs will not corroborate epidemiologic findings.

Testing of these hypotheses is a useful approach for investigating the potential risks from exposure to mixtures of DBPs; studies can be designed to address these hypotheses. Although it may be noted that these hypotheses are not mutually exclusive, the objective of this paper is to examine how scientists can address the first hypothesis, that adverse health effects exist and are attributable to exposure to the complex mixture. Furthermore it is postulated that high-dose, single chemical animal toxicity studies alone are inadequate for characterizing DBP health risks and that a concentrated effort by a multidisciplinary team of researchers is necessary to generate the appropriate data for this problem.

Producing Appropriate Data for Risk Assessment

Several risk assessment issues are of concern to managers responsible for ensuring safe drinking water for the public. The first issue, as indicated above, is to establish or refute the actual existence of human health risks from exposure to DBPs. Because the answer to this is not imminent, some drinking water regulations have been promulgated and others posed with the goal of controlling levels of DBPs in the drinking water (U.S. EPA, 1979, 1994, 1997, 1998b). As rules go into effect, alternative drinking water treatment technologies are developed to meet these new standards. Thus, to choose among treatment options, it is necessary to evaluate potential changes in health risks across various drinking water treatment systems and source waters.

A recent report assessing the toxicity of DBP mixtures was prepared by the International Risk Science Institute (ILSI, 1998) from recommendations by an expert panel. The report specifies that DBP risks cannot be assessed using single chemical information alone. The panel suggested that mixtures testing focus on three scenarios: defined or simple mixtures of DBPs; reproducible disinfection scenarios (RDS) in which water samples are produced by controlling the characteristics of source water and then subjecting it to specific treatment trains; and real world drinking water samples or their extracts. In addition, researchers were presented with a

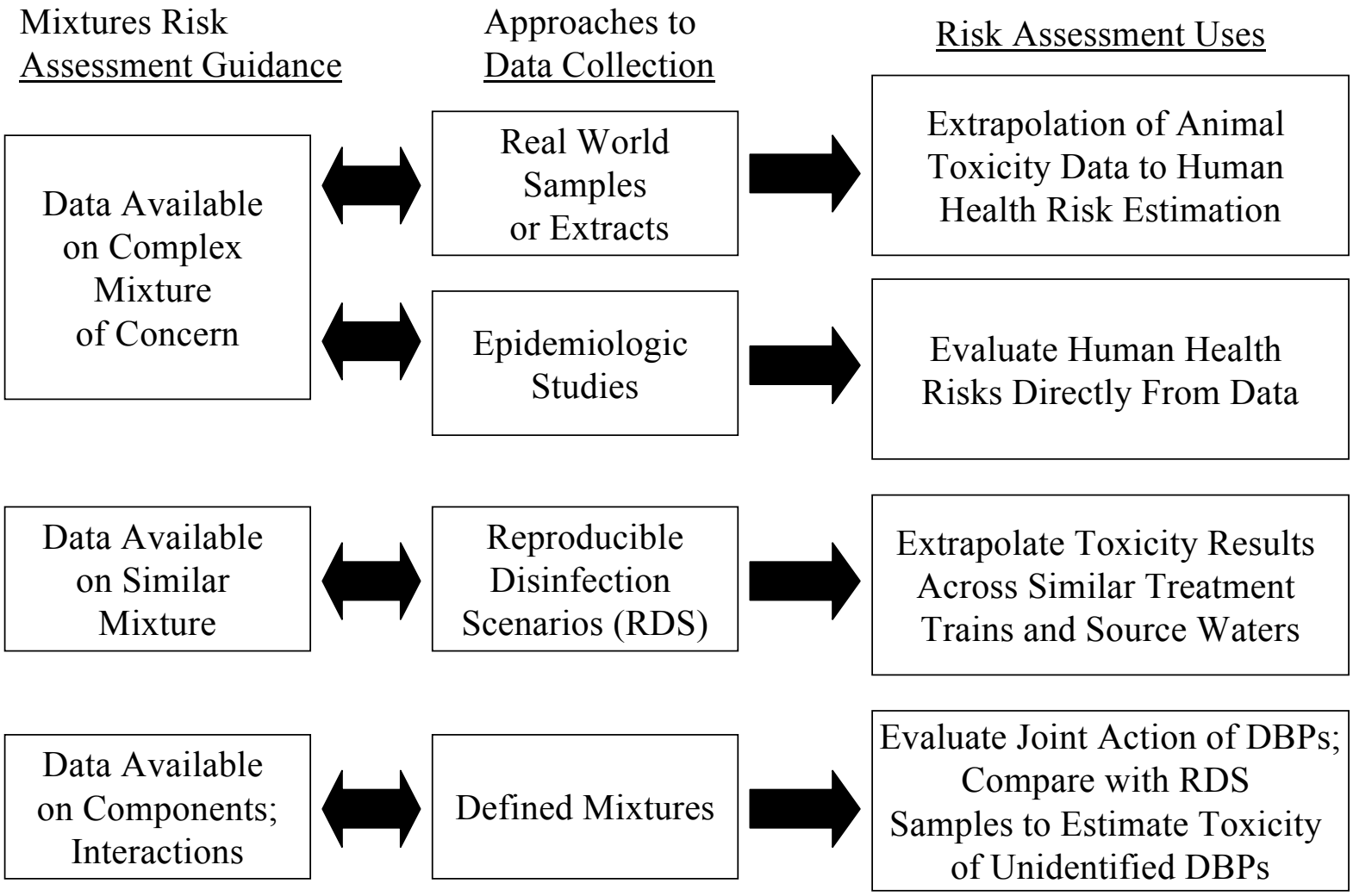
3-tiered testing approach for mixtures from simple (*in vitro* tests) to short-term screening (medaka assays, 1-90 day *in vivo* studies) to more sophisticated studies (*in vivo* chronic bioassays; epidemiologic studies). It may be inferred from this report that a blend of the three testing scenarios and the 3-tiered testing approach could be used to characterize DBP mixtures toxicity.

The U.S. EPA (1986) published *Guidelines for the Health Risk Assessment of Chemical Mixtures*, in which three approaches to quantifying health risk for a chemical mixture are recommended depending on the type of data available to the risk assessor: data on the complex mixture of concern; data on a similar mixture; or data on the individual components of the mixture or on their interactions. Figure 1 illustrates that these three scenarios can be mapped directly to the ILSI testing scenarios mentioned above and shows the potential uses of such data for risk assessment. In the first approach, toxicity data are available on the complex mixture of concern, in this case, real world drinking water samples or extracts. The quantitative risk assessment is done directly from these preferred data, which can include either epidemiologic or toxicologic data. Although there are advantages to testing the complex mixture of concern, this can also be problematic. Toxicologic data from animal studies suffer from the uncertainties inherent in extrapolation to human health risk. Epidemiologic data appear to be superior in terms of evaluating humans directly, but these data suffer from confounding sources of potential toxicity that are difficult to account for.

The second approach uses data on a "sufficiently similar" mixture, which is represented for DBPs by testing the RDS samples. RDS samples may be classified as similar mixtures whose toxicity data may provide a measure of the expected toxicity of finished drinking water for a given treatment train and source water characterization. Data in an RDS toxicity database could be judged for similarity against data from a treatment plant by identifying and measuring the concentrations of DBPs in each, comparing component proportions across the mixtures, and contrasting available toxicity data on their components. For the RDS mixtures, the source of the mixtures is controlled and measured so comparisons could be made across similar treatment trains and source waters. Analytical chemistry techniques would characterize the composition and stability of the mixtures. Available toxicity data on the components could be used to characterize expected health effects. If the treatment plant data and the RDS mixture are judged to be similar, then the quantitative risk assessment for the treatment process may be derived from health effects data on RDS sample.

Finally, the third mixtures risk assessment approach that maps to testing defined mixtures of DBPs is to evaluate the mixture through an analysis of its components. For example, one could use dose addition for risk estimation of systemic effects and response addition for estimates of cancer risk. These particular procedures include a general assumption that interaction effects at low concentration levels either do not occur at all or are small enough to be insignificant to the risk characterization. Other, newer approaches are under development that incorporate interactions information into these risk estimates when such data are available (Hertzberg et al., 1999; U.S. EPA, 1999). For DBPs, toxicity and concentration data on the components of a mixture can be combined and added (depending on the assumption used) to

Figure 1
Mapping of Risk Assessment Approaches to Drinking Water Studies



estimate mixtures risk. Again, analytical chemistry techniques are needed to identify and quantify the components.

For the toxicologic data, each of these three approaches to data collection can be subject to the 3-tier testing scheme mentioned above. Because more than several hundred potential DBPs have been identified, simple test systems can be used to find the most toxic combinations of chemicals for testing in more sophisticated and expensive assays. Comparisons of results among the three approaches to data collection will also be valuable. For example, the results from real world samples can be compared with those of defined mixtures to determine how much toxicity may be produced by DBPs that are not individually identified.

Testing Procedures for DBP Mixtures

Under the three mixtures/three tiers scheme suggested above (ILSI, 1998), a logical series of toxicity investigations can be designed to target the risk assessment questions of interest. Furthermore, experiments can be carefully designed by an interdisciplinary team of scientists so the resulting data can be used to validate quantitative risk assessment statistical methods, to characterize the toxic behavior of DBP mixtures, and to make inferences about DBP health risk. Such a multidisciplinary team would likely include water treatment engineers, laboratory scientists, toxicologists, epidemiologists, analytical chemists, statisticians, and risk assessment scientists.

Drinking Water Samples

Studies performed directly on drinking water samples, particularly epidemiologic studies, can provide health data on complex mixtures that represent finished drinking water. Epidemiologic studies have been useful for the hazard identification of potential health effects. When appropriate data are available, quantitative risk estimates, such as relative risk ratios, may also be produced. For the chlorinated drinking water studies to date, there have been various methods of exposure characterization that include comparing persons living in areas supplied with either chlorinated water or surface water with persons living in areas supplied with water described as nonchlorinated, chloraminated, or ground water. Such broad categories of exposure classification limit the usefulness of the results to making broad statements about potential health effects from chlorination. They are not useful for evaluating risks across specific treatment options and source waters. Attributable risk (AR) estimates are desirable and can be defined as the "proportion of the cases of disease occurring among exposed persons which is in excess in comparison with the non-exposed" (Breslow and Day, 1980). These AR calculations include an assumption of causality, and although epidemiologic data has weakly associated chlorinated water with health effects, causality between these has not yet been established. Thus, epidemiologic studies logically provide the best data for risk assessment, but such studies must be well designed to account for confounding factors, to accurately assess exposure, to address issues of causality, and to answer specific risk assessment questions.

Because chronic *in vivo* exposure of animals directly to drinking water samples or their extracts at environmentally relevant concentrations has shown no effects, continuation of these

studies may provide little scientific gain. It can only be assumed that if effects in humans are real, then differences exist that preclude the appearance of adverse effects in animals, such as differences in mechanism of action, duration of exposure, or multi-route effects. On the other hand, the use of *in vitro* studies to screen for mutagenicity or short-term *in vivo* screening studies, such as 96-hour testing for developmental effects in Japanese medaka (*Oryzias latipes*) fish embryos, could be useful to look for tier I differences in effects for drinking water samples.

Defined Mixtures

Testing of defined mixtures of DBPs is necessary to investigate whether interaction effects (greater than or less than additive) among chemicals exist at low human exposures and, if so, to estimate their magnitude and direction. In the absence of interaction effects, it may be possible to establish additivity as a valid assumption for risk assessment of DBPs, as this affects the method selected for characterizing mixtures risk. There is precedent for the use of response addition (Gaylor et al., 1997; U.S. EPA, 1989b) and dose addition (U.S. EPA, 1989a,b) in the evaluation of mixtures risk as well as through other definitions of additivity (Chen et al., 1989; Gennings et al., 1997; U.S. EPA, 1998c). These approaches are all component-based and do not include interactions information because they are generally applied to mixtures that occur together at low concentration levels. Component-based approaches are useful in the comparison of DBP risks across different drinking water treatment systems because for each system, specific DBPs can be identified and their concentrations measured (U.S. EPA, 1998a,c). The issue is actually more complex than indicated here because of variations in DBP production resulting from many factors, including source water characteristics, seasonal changes, pH levels, etc. However, the validation of additivity through laboratory experimentation as a meaningful assumption for estimation of DBP risk under specific conditions is still important to establish.

Defined mixtures testing can also help determine what is happening at lower doses of DBPs than can be used in single chemical studies. In addition, precursor effects for more adverse effects such as cancer can be identified. For example, Gennings et al. (1997) showed evidence of additivity for the liver enzyme, sorbitol dehydrogenase (SDH), when mixtures of the four THMs were tested in CD-1 mice. Because it has recently been suggested that chloroform is a non-genotoxic carcinogen in the liver (U.S. EPA, 1998b), evidence of a possible precursor effect at relatively low doses of this four chemical mixture is important. There is limited data on defined mixtures studies, and what is available is mostly on the THMs (see U.S. EPA, 1998c). Investigations of defined mixtures for other chemical classes and across classes, e.g., mixtures of THMs, HAAs, bromate, and other DBPs, is important for increased understanding of DBP mixtures toxicity.

Reproducible Disinfection Scenarios

In addition to the investigation of defined mixtures, it is important to also study samples developed using reproducible disinfection scenarios (RDS). Studies using such samples can provide toxicity data on complex mixtures that represent various combinations of source water characteristics and drinking water treatment trains for health risk estimation using a similar mixtures approach to risk assessment. Many DBPs in drinking water have yet to be identified,

and of those that are known, toxicity data are available for relatively few. To characterize the toxicity of the unidentified halo-organic and inorganic materials that exist in finished drinking water, studies must be performed on both defined mixtures and on complex mixtures (RDS samples). Toxicity studies on RDS samples can be performed in conjunction with studies on defined mixtures and the resulting toxicities compared. In this way, the toxicity of the RDS samples can be separated into portions attributable to the defined components (as estimated from the defined mixture results) and portions attributable to the unidentified materials.

As a simplified example, suppose we perform a short-term rat study using a control group and two dose groups, one to be dosed with a simple defined mixture of the four THMs and another to be dosed with an amount of RDS sample that is THM-equivalent to the defined THM dose. Furthermore, suppose there is an increase in the liver enzyme, sorbitol dehydrogenase (SDH) of 15% and 30% above the control group for the defined mixture and the RDS sample, respectively. Then, the additional 15% response can be attributed to the portion of the complex mixture that is not THM material. Such data that can provide evidence on the toxicity of unidentified material can be used in health risk assessment.

An RDS approach, although conceptually appealing, may be difficult to execute, primarily because the integrity of the samples must be assured prior to use in toxicity experiments. In addition, statistical methods must be developed to estimate the toxicity of the unidentifed DBP. A proposed plan for such an effort entails the following six components.

1) Production of RDS samples by drinking water engineers. These would undergo toxicity testing after appropriate concentration and storage procedures. Integrity of the samples must be assured so the mixture tested fairly represents the conditions under which the samples are produced.

2) Analyses of samples by analytical chemists to determine the amounts and proportions of both known and unknown DBPs in the mixture and to test for the integrity of the mixture.

3) Experimental design by a team of laboratory scientists, statisticians, and risk assessment scientists to maximize the usefulness of the toxicity data that are generated.

4) Toxicity tests by laboratory scientists. Doses are calibrated by using an amount of the mixture corresponding to amount(s) of known single DBPs to determine how much of the observed toxicity is attributable to the known and unknown chemicals. Testing should follow the three tiers suggested in the ILSI (1998) report and include all the relevant health effects: liver, kidney, bladder, colon, and rectal cancer; mutagenicity, nephrotoxicity, and hepatotoxicity; neurotoxicity; and developmental and reproductive endpoints. Testing should be performed for dermal, oral, and inhalation routes.

5) Development of Physiologically Based Pharmacokinetic (PBPK) models, other biologically based models, or statistical models to explain the toxic behavior of DBP mixtures at low levels, at which effects are not observable in animal studies.

6) Use of data for human risk estimation by risk assessment scientists to perform actual risk characterizations and to validate new risk assessment methods.

Conclusions

High-dose, single-chemical, animal studies are useful for many environmental problems such as the evaluation of pesticide toxicity or setting limits on certain chemicals produced by municipal waste combustion. They are useful in determining the mechanism of action for a chemical's toxicity and the establishment of toxicity thresholds. These studies are also highly applicable to chemicals for which human exposures are to high doses. In the case of disinfection by-products produced by drinking water disinfection however, these studies alone are inadequate to characterize risk. It is key to this risk assessment problem that there is a highly variable complex mixture of chemicals produced, and that the human exposure is to extremely low doses of DBPs over time. Thus, a series of well planned testing procedures should be performed that are designed to address specific research needs and that are coordinated among concerned scientists. These should include the three approaches to data collection, the 3-tiered testing scheme, consideration of both epidemiologic and toxicologic testing, and studies that are designed to address risk assessment issues.

References

Bove, F.J., M.C. Fulcomer, J.B. Klotz, J. Esmart, E.M. Dufficy, J.E. Savrin. 1995. Public drinking water contamination and birth outcomes. *Am. J. Epidemiol.* 141(9):850-862.

Breslow NE, Day NE. Statistical Methods in Cancer Research, Vol. 1 - the Analysis of Case-Control Studies. Iarc Scientific Publications No. 32;1980:73-81.

Bull, R.J. and F.C. Kopfler. 1991. Health Effects of Disinfectants and Disinfection By-Products. AWWA Research Foundation.

Bull, R.J., Robinson, M., Meier, J.R., Stober, J. 1982. Use of biological assay systems to assess the relative carcinogenic hazards of disinfection by-products. *Environ. Health Perspec*. 46:215-227.

Cantor, K.P., R. Hoover, P. Hartge et al. 1985. Drinking water source and bladder cancer: A case-control study. In: Water Chlorination: Chemistry, Environmental Impact and Health Effects, Vol. 5. Jolley, R.L., R.J. Bull and W.P. Davis et al., Ed. *Lewis Publishers, Inc., Chelsea, MI.*
p. 145-152.

Cantor, K.P., C.F. Lunch, F. Hildesheim et al. 1997. Drinking water source and chlorination byproducts. I. Risk of bladder cancer. *Epidemiol.* 9:21-28.

Chen, JJ, Heflich, RH, Hass, BS. 1989. A Response-Additive Model for Analyzing Mixtures of Two Chemicals in the Salmonella Reversion Assay. *Biomed. J.* 31:495-503.

Freedman, M., K.P. Cantor, N.L. Lee et al. 1997. Bladder cancer and drinking water: A population-based case-control study in Washington County, Maryland (United States). *Cancer Causes and Control*. 8:738-744.

Gaylor, D.W., Axelrad, J.A., Brown, R.P., Cavagnaro, J.A., Cyr, W.H., Hulebak, K.L., Lorentzen, R.J., Miller, M.A., Mulligan, L.T., and Schwetz, B.A. 1997. Health Risk Assessment Practices in the U.S. Food and Drug Administration. *Regulatory Toxicology and Pharmacology*. 26:307-321.

Gennings, C, Schwartz, P, Carter, WH Jr., Simmons, JE. 1997. Detection of Departures from Additivity in Mixtures of Many Chemicals with a Threshold Model. *J. Agric. Biol. Environ. Stat.* 2:198-211.

Hertzberg, R.C, G. Rice, L.K. Teuschler, 1999. Methods for Health Risk Assessment of Combustion Mixtures. In: Hazardous Waste Incineration: Evaluating the Human Health and Environmental Risks. S. Roberts, C. Teaf, and J. Bean eds. CRC Press LLC. pp. 105-148.

ILSI. 1998. Assessing to Toxicity of Exposure to Mixtures of Disinfection By-Products Research Recommendations. ILSI Risk Science Institute, Washington, DC.

Jacangelo, J.G., N.L. Patania, K.M. Reagan et al. 1989. Ozonation: Assessing its role in the formation and control of disinfection byproducts. *JAWWA*. 81:74-84.

Kavlock, R., Chernoff, N., Carver, B., Kopfler, F. 1979. Teratology studies in mice exposed to municipal drinking water concentrates during organogenesis. *Fd. Cosmet. Toxicol*. 17:343-347.

King, W.D., L.D. Marrett. 1996. Case-control study of water source and bladder cancer. *Cancer Causes and Control*. 7:596-604.

Kool, H.J., van Kreijl, C.F., van Kranen, H.J. 1981. The use of XAD-resins for the detection of mutagenic activity in water. II. Studies with drinking water. *Chemosphere* 10:99-108.

Kramer, M.D., C.F. Lynch, P. Isacson, J.W. Hanson. 1992. The association of waterborne chloroform with intrauterine growth retardation. *Epidemiol*. 3(5):407-413.

Krasner, S.W., M.J. McGuire, J.G. Jacangelo et al. 1989. The occurrence of disinfection byproducts in U.S. drinking water. *JAWWA*. 81:41-53.

Loper, J.C., Lang, D.R., Schoeny, R.S., Richmond, B.B., Gallagher, P.M., Smith, C.C. 1978. Residue organic mixtures from drinking water show *in vitro* mutagenic and transforming activity. *J. Toxicol. Environ. Health* 4:919-938.

Lykins Jr., B.W., W.E. Koffskey, K.S. Patterson. 1994. Alternative disinfectants for drinking water treatment. *J. Environ. Engin.* 120(4):745-758.

McGeehin, M.A. et al. 1993. Case-control study of bladder cancer and water disinfection methods in Colorado. *Am. J. Epidemiol.* 138:492-501.

Miltner, R.J., E.W. Rice, A.A. Stevens. 1990. Pilot-Scale Investigation of the Formation and Control of Disinfection Byproducts. In: 1990 Annual Conference Proceedings. AWWA Annual Conference, Cincinnati, Ohio. 2:1787-1802.

Nestman, E.R., Otson, R., LeBel, G.L., Williams, D.T., Lee, E.G.H., Biggs, D.C. 1982. Correlation of water quality parameters with mutagenicity of drinking water extracts. In: Water Chlorination: Environmental Impact and Health Effects, Vol. 4. Eds. R.L. Jolley, W.A. Brungs and R.B. Cumming. Ann Arbor Science Publ., Ann Arbor, MI.

NTP. (National Toxicology Program) 1985. Toxicology and carcinogenesis studies of chlorodibromomethane in F344/N rats and B6C3F mice (gavage studies). NTP TR282.

NTP. (National Toxicology Program) 1986. Toxicology and carcinogenesis studies of bromodichloromethane in F344/N rats and B6C3F mice (gavage studies). NTP Technical Report, Ser. No. 321, NIH Publ. No. 87-2537.

NTP. (National Toxicology Program) 1989. Toxicology and carcinogenesis studies of tribromomethane and bromoform in F344/N rats and B6C3F mice (gavage studies). NTP-350.

Pereira, M.A. and Phelps, J.B. 1997. Promotion by mixtures of dichloroacetic acid and trichloroacetic acid of N-methyl-N-nitrosourea-initiated cancer in the liver of female B6C3F1 mice. Cancer Letters 115:15-23.

Richardson, S.D. 1998. Identification of Drinking Water Disinfection By-Products. In: John Wiley's Encyclopedia of Environmental analysis and Remediation. Robert A. Meyers, Ed. 3:1398-1421.

Smith, M.K., J.L. Randall, E.J. Read and J.A. Stober. 1989. Teratogenic activity of trichloroacetic acid in the rat. *Teratology*. 40:445-451.

Swan, S.H., K. Waller, B. Hopkins et al. 1998. A prospective study of spontaneous abortion: Relation to amount and source of drinking water consumed in early pregnancy. *Epidemiol.* 9(2):126-133.

U.S. EPA. 1979. National Interim Primary Drinking Water Regulations; Control of Trihalomethanes in Drinking Water. FR 44(231):68624-68707.

U.S. EPA. 1986. Guidelines for the Health Risk Assessment of Chemical Mixtures. Federal Register.51(185):34014-34025.

U.S. EPA. 1989a. Interim Procedures for Estimating Risks Associated with Exposures to Mixtures of Chlorinated Dibenzo-p-dioxins and -dibenzofurans (CDDs and CDFs) and 1989 update. EPA/625/3-89/016. Risk Assessment Forum. March.

U.S. EPA. 1989b. Risk Assessment Guidance for Superfund. Vol. 1. Human Health Evaluation Manual (Part A). EPA/540/1-89/002.

U.S. EPA. 1998b. National Interim Primary Drinking Water Regulations; Disinfectants and Disinfection Byproducts; Notice of Data Availability; Proposed Rule. FR 63(61):15674-15692.

U.S. EPA. 1998c. Workshop Report: The Risk Assessment of Mixtures of Disinfection By-Products (DBPs) in Drinking Water. NCEA-C-0459.

U.S. EPA. 1999. *Guidance for Conducting Health Risk Assessment of Chemical Mixtures*. External Review Draft. NCEA-C-0148.

Waller, K., S.H. Swan, G. DeLorenze, B. Hopkins. 1998. Trihalomethanes in drinking water and spontaneous abortion. *Epidemiol*. 9(2):134-140.

A MULTIPLE-PURPOSE DESIGN APPROACH TO THE EVALUATION OF RISKS FROM MIXTURES OF DISINFECTION BY-PRODUCTS*†

Linda K. Teuschler[1], Chris Gennings[2], William M. Stiteler[3], Richard C. Hertzberg[1], Joan T. Colman[3], Arunthavarani Thiyagarajah[4], John C. Lipscomb[1], William R. Hartley[4], Jane Ellen Simmons[5].

[1]NCEA, U.S. EPA, Cincinnati, OH,
[2]Virginia Commonwealth University, Richmond, VA,
[3]Syracuse Research Corporation, Syracuse, NY,
[4]Tulane University, Environmental Diseases Prevention Research Center, New Orleans, LA,
[5]NHEERL, U.S. EPA, Research Triangle Park, NC

June 1999

**Citation for this work: Teuschler, L.K., C. Gennings, W.M. Stiteler, R.C. Hertzberg, J.T. Colman, A. Thiyagarajah, J.C. Lipscomb, W.R. Hartley, J.E. Simmons. 2000. A Multiple-Purpose Design Approach to the Evaluation of Risks from Complex Mixtures of Disinfection By-Products (DBPs). Drug and Chemical Toxicology. 23(1):307-321.*

†Also see errata following this section

Disclaimer

The research described in this article has been reviewed by the National Health and Environmental Effects Research Laboratory and the National Center for Environmental Assessment - Cincinnati Office, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Abstract

Drinking water disinfection has effectively eliminated much of the morbidity and mortality associated with waterborne infectious diseases in the United States. Various disinfection processes however, produce certain types and amounts of disinfection by-products (DBPs), including trihalomethanes (THM), haloacetic acids, haloacetonitriles, and bromate, among others. Human health risks from the ubiquitous exposure to complex mixtures of DBPs are of concern because existing epidemiologic and toxicologic studies suggest the existence of systemic or carcinogenic effects. Researchers from several organizations have developed a multiple-purpose design approach to this problem that combines efficient laboratory experimental designs with statistical models to provide data on critical research issues (e.g., estimation of human health risk from low-level DBP exposures, evaluation of additivity assumptions as useful for risk characterization, estimation of health risks from different drinking water treatment options). A series of THM experiments have been designed to study embryonic development, mortality and cancer in Japanese medaka (*Oryzias latipes*) and liver and kidney endpoints in female CD-1 mice. The studies are to provide dose-response data for specific mixtures of the four THMs, for the single chemicals, and for binary combinations. The dose-levels and mixing ratios for these experiments were selected to be useful for development and refinement of three different statistical methods: testing for departures from dose-additivity; development of an interactions-based hazard index; and use of proportional-response addition as a risk characterization method. Preliminary results suggest that dose-additivity is a reasonable risk assessment assumption for DBPs. The future of mixtures research will depend on such collaborative efforts that maximize the use of resources and focus on issues of high relevance to the risk assessment of human health.

Introduction

Since the early part of the 20th century, drinking water disinfection has effectively eliminated much of the morbidity and mortality associated with waterborne infectious diseases in the United States. Disinfection processes, however, produce a large number of disinfection by-products (DBPs) that vary in the type and amount produced depending on the disinfectant used (e.g., chlorine, ozone, chlorine dioxide or chloramine) and on source water characteristics. The DBPs that have been identified and on which some toxicity and water concentration data are available include the trihalomethanes (THMs), haloacetic acids (HAAs), haloacetonitriles, haloketones, aldehydes, bromate, chloral hydrate, and chloropicrin, among others[2, 1, 3, 4]. More recently, Richardson (1998)[5] identified several hundred potential DBPs from various disinfection scenarios. Of the identified DBPs, less than 20 are well studied toxicologically.

Human health risks from ubiquitous low-level exposure to complex mixtures of DBPs are of concern because existing epidemiologic and toxicologic studies suggest the possibility of systemic or carcinogenic effects. Toxic effects from exposure to low-concentration environmental mixtures have not been found in animal studies[6, 7]. Epidemiologic studies, on the other hand, of chlorinated drinking water exposures in humans suggest weak associations with bladder, rectal and colon cancer[8, 9, 10, 11, 12] and limited evidence of reproductive and developmental effects[13, 14, 15, 16]. There are a few studies available on defined mixtures of DBPs that show evidence of dose-additivity for hepatotoxicity in mice exposed to mixtures of THMs[17] and of synergistic activity by mixtures of HAAs for promotion of cancer[18]. Most DBP toxicity data is from single chemical *in vivo* or *in vitro* studies. There is evidence in single-chemical animal studies at high DBP dose levels of reproductive, developmental, hepatotoxic and nephrotoxic effects and carcinogenicity[19, 20, 21, 22, 23]. Finally, there is evidence of mutagenicity from exposure to extracts of finished drinking water in *in vitro* studies[24, 25, 26].

Humans are chronically exposed to low concentrations of variable mixtures of DBPs. Exposure occurs via drinking water as well as by dermal and inhalation routes from sources such as indoor air and shower water[27]. Therefore, laboratory investigations using efficient experimental designs are needed to examine the effects that are produced by exposure to DBP mixtures. Furthermore, development of risk characterization methods is essential for estimation of human health risks. Mixtures studies, however, are traditionally thought to be complicated to perform and prohibitively expensive, yielding results that are difficult to interpret. Full factorial designs (see, for example,[28]) provide good data on interaction effects among chemicals, but are limited by practical constraint to a few chemicals (c) and dose levels (m) because the number of required dose groups is then large (m^c). Fractionated factorial designs for laboratory studies reduce this restriction[29], but these can also be resource intensive. Thus, it is important to develop methods that offer efficient experimental designs for the laboratory that will also produce easily interpretable results. Statistical methods have been designed to detect departures from additivity using smaller experiments[30, 31, 32]. The resulting toxicity data can then be used either directly in a DBP mixtures risk characterization or to support the underlying assumptions of risk assessment methods that are under development.

In this paper, researchers from several organizations have developed a multiple-purpose design approach to investigating the toxicity of DBP mixtures. The goal is to provide data and methods relevant to critical research issues, such as: the estimation of human health risk from low-level multi-chemical DBP exposures; assessment of various additivity assumptions as useful defaults for risk characterization; and calculation of health risk estimates for different drinking water treatment options. Areas of collaboration for this multi-disciplinary team of investigators include efficient experimental design development, toxicologic evaluation of single DBPs and mixtures using mice (CD-1) and fish (Japanese medaka - *Oryzias latipes*), comparison of chlorination and ozonation mixing ratios, and use of data to develop and refine risk assessment methods.

Multiple-Purpose Design Approach to DBP Mixtures Research

The multiple-purpose design approach involves the collaboration of investigators from two experimental laboratories and three groups interested in environmental statistics and mixtures risk assessment. The project areas and lead organizations are as follows:

- Investigators from the U.S. EPA's National Health and Environmental Effects Research Laboratory (NHEERL) are assessing liver and kidney toxicity in female CD-1 mice following exposure by oral gavage in an aqueous vehicle for 14 days.

- Researchers at Tulane University are developing a medaka embryo exposure methodology to more efficiently screen single DBPs and mixtures to assess developmental and chronic (cancer, reproductive effects and other systemic endpoints) toxicity.

- Statisticians at Virginia Commonwealth University are investigating efficient experimental designs, including the development of statistical threshold models for mixtures built from single chemical data to test for departures from dose-additivity for THM mixtures.

- Risk assessors at Syracuse Research Corporation and the U.S. EPA's National Center for Environmental Assessment - Cincinnati Office are developing two additivity-based risk assessment methods, an interaction-based hazard index and proportional-response addition, as the basis for expressing DBP mixtures risk. These assessments will utilize the data generated by NHEERL and Tulane.

Experimental Data

To illustrate the multiple-purpose design approach, the balance of this paper will show data on liver endpoints in female CD-1 mice. The studies provide dose-response data for specific mixtures of the 4 THMs, which are chloroform ($CHCl_3$), bromoform ($CHBr_3$), chlorodibromomethane (CDBM), and bromodichloromethane (BDCM), as well as for the single chemicals and all possible binary combinations.

The research objectives are to: develop an efficient experimental design for the collection of data on mixtures; provide data for the development of the threshold additivity model; produce data useful in testing the proportional-response addition and interaction-based hazard index risk assessment methods; and develop an understanding of the toxicity (i.e., the potency and nature of the interaction) of the 4 THMs. The experimental animal was selected to explore the possibility of using single chemical data from the published literature to construct the expected response of a chemical mixture rather than repeatedly generating the single chemical curves for every mixtures experiment. Female CD-1 mice were selected as only one report[33] was found in the published literature in which the hepatotoxicity of each of the four THMs was investigated in the same laboratory following subacute administration in an aqueous vehicle. The liver and kidney were identified as critical target organs for the THMs based on EPA-verified chronic oral reference doses (RfDs)[34].

In the female CD-1 mice experiments, groups of 8-20 animals, 65-70 days old at the initiation of exposure, were administered THMs by oral gavage for 14 days in an aqueous vehicle (10% Alkamulus EL-620) at a constant volume of 10 ml/kg. Dosing solutions were made daily. Mice were dosed daily for 14 days and were sacrificed on Day 15. At sacrifice, several serum endpoints and body and liver weights were evaluated. The percent relative liver weight was calculated as the liver to body weight ratio times one-hundred. Serum was analyzed for indicators of hepatic toxicity: sorbitol dehydrogenase (SDH), alanine aminotransferase (ALT), and aspartate aminotransferase (AST). Histopathologic analysis was conducted with centrilobular necrosis as a primary finding. The centrilobular necrosis was characterized as none (no necrotic hepatocytes seen), mild (less than 10% of the hepatocytes involved), mild-moderate (less than 50% of the hepatocytes involved), moderate (approximately 50% of the hepatocytes involved), moderate -severe (more than 50% of the hepatocytes involved), and severe (more than 75% of the hepatocytes involved). Table I shows an example of the experimental results for these endpoints from a single chemical experiment with chloroform. Similar results were found for all 4 THMs, showing dose-related increases in relative liver weight, serum enzyme levels and incidence and severity of centrilobular necrosis.

Three types of experiments were designed to study: 1) various 4-THM mixtures; 2) the single THMs; and 3) each of the six binary combinations. To assess the toxicity of the single chemicals, each THM was assessed in a separate experiment. The following dosages were given for each of the single THMs: 0, 0.152, 0.305, 0.76, 1.52, or 3.05 mmol/kg/d (Table I). Four mixture experiments (of 2 total dosages each) were also conducted for combinations of the 4 THMs as shown in Table II; concurrently, single-chemical data points were included to provide for comparison with single chemical dose-response curves. In one experiment, a prototype mixture, based on the work of Krasner et al. (1989)[1], with the mixing ratio of the THMs based on

Table I Example Hepatotoxicity Data for Chloroform						
Mice N	Dose mmol/kg/day	ALT μ (σ) IU/l	AST μ (σ) IU/l	SDH μ (σ) IU/l	% Relative Liver Wgt μ (σ)	Centrilobular Necrosis Results* (Number of animals)
12	0	18.5 (4.1)	42.6 (5.9)	15.6 (3.27)	5.1 (0.57)	None (11) Mild (1)
13	0.152	15.1 (4.5)	41.6 (7.5)	16.8 (4.2)	4.64 (0.54)	None (13)
15	0.305	21.2 (7.0)	47.3 (18.0)	21.3 (5.85)	5.04 (0.39)	None (7) Mild (4) Mild-moderate (3) Moderate (1)
15	0.76	50.0 (83.0)	65.7 (65.0)	30.3 (19.9)	5.28 (0.63)	None (1) Mild (1) Mild-moderate (1) Moderate (8) Moderate-severe (2) Severe (2)
13	1.52	80.2 (81.0)	84.8 (53.0)	50.8 (22.5)	6.28 (0.45)	Mild (1) Moderate (7) Moderate-severe (3) Severe (2)
12	3.05	125.0 (64.0)	117.0 (42.0)	80.2 (9.35)	6.82 (0.70)	Moderate (4) Moderate-Severe (2) Severe (6)

*Key: Centrilobular necrosis was characterized as none (no necrotic hepatocytes seen), mild (less than 10% of the hepatocytes involved), mild-moderate (less than 50% of the hepatocytes involved), moderate (approximately 50% of the hepatocytes involved), moderate -severe (more than 50% of the hepatocytes involved), and severe (more than 75% of the hepatocytes involved).

Table II
Mixture Doses of THMs Tested in Female CD-1 Mice
(mmoles/kg/day)

THM	Mix #1	Mix #2	Mix #3	Mix #4	Mix #5	Mix #6	Mix #7	Mix #8
BDCM	0.208	0.104	0.171	0.0935	0.513	0.2805	1.026	0.561
CDBM	0.084	0.042	0.145	0.0645	0.435	0.1935	0.87	0.387
$CHC1_3$	0.568	0.284	0.1595	0.3275	0.4785	0.9825	0.957	1.965
$CHBr_3$	0.012	0.0006	0.0245	0.0145	0.0735	0.0435	0.147	0.087
Total Dosage	0.872	0.436	0.5	0.5	1.5	1.5	3.0	3.0

Note: Mix #1, 2 are average ratios for 35 treatment facilities.[1]
Mix #3, 5, 7 are ratios from a chlorination process.
Mix #4, 6, 8 are ratios from an ozonation process followed by chloramination.

the average seasonal proportions at 35 water treatment facilities, was tested at two mixture dosages (Mix #1,2). Three additional experiments were conducted using mixing ratios of the THMs for two water treatment processes, chlorination (Mix #3,5,7) and ozonation followed by chloramination (Mix #4,6,8).

Six binary experiments are planned and underway, each testing a binary mixture (A, B) of the four THMs. Each experiment is to consist of 12 treatment groups: one vehicle control group, three dose levels of chemical "A" alone and three dose levels of chemical "B" alone, and five mixture groups. Table III illustrates this design for the binary combination of $CHCl_3$ and BDCM. Three of the five mixture groups are at a 1:1 ratio of chemicals A:B with varying total dosages (0.10, 1.0, or 3.0 mmol/kg/d). The mixing ratio of the other two mixture groups is based on the proportions of the two chemicals found in 'typical' treated drinking water[1] and will vary depending on the pair of chemicals considered. The total dosages of these two mixture groups will be held constant at 3.0 and 1.0 mmol/kg/d. In Table III, for example, the binary mixture of $CHCl_3$:BDCM is to be tested at a ratio of 2.7:1.0. The effect of the mixing ratio on observed interactions can be evaluated by the use of two different mixing ratios in each experiment. Finally, the use of mixtures with ratios found in treated drinking water allows for examination of the likelihood of nonadditive interactions at ratios that have environmental relevance.

Statistical Methods

The dose-levels and mixing ratios described above for these experiments were selected to be useful for development of three different statistical methods: testing of departures from dose-additivity at low doses with a threshold model; estimation of mixtures toxicity from single chemical and binary data using an interaction-based hazard index; and use of proportional-response addition as a risk characterization method. Expanded descriptions for these methods are available in a 1998 EPA workshop report[35] and in the cited references below.

Statistical Testing of Departures from Additivity

Methodology for testing of departures from additivity can be used to help define the joint action of defined mixtures of chemicals[32, 17]. It is a useful method to either establish or refute the use of dose-additivity as an assumption in mixtures risk assessment. This procedure uses single chemical dose-response data to build a joint dose-response model under the assumption that non-additive interactions are not occurring among the chemicals in the mixture. Then, laboratory data on the mixture can be compared to the model prediction(s) using a prediction interval(s) to determine if the joint action can be classified as additive or characterized as greater than additive (synergism) or less than additive (antagonism). The continuous model used for liver enzyme data in CD-1 mice is shown in equation (1)[17].

$$\mu_{ij}^{\lambda} = \begin{cases} \beta_0 & \text{if } \sum\beta_i x_{ij} < \delta \\ \beta_0 + \sum\beta_i x_{ij} - \delta & \text{if } \sum\beta_i x_{ij} \geq \delta \end{cases}, \qquad (1)$$

Table III Binary Experimental Design for CHC_{l3} and BDCM	
Chemical	Dose group (mmol/kg/d)
Control (vehicle only)	0
$CHCl_3$ alone	0.1 1.0 3.0
BDCM alone	0.1 1.0 3.0
1:1 $CHCl_3$:BDCM	0.1 (0.05 $CHCl_3$:0.05 BDCM) 1.0 (0.5 $CHCl_3$:0.5 BDCM) 3.0 (1.5 $CHCl_3$:1.5 BDCM)
2.7:1 $CHCl_3$:BDCM	1.0 (0.73 $CHCl_3$:0.27 BDCM) 3.0 (2.2 $CHCl_3$:0.8 BDCM)

where:

μ_{ij}^{λ} = continuous endpoint response for the mixture (where λ is a pre-specified transformation parameter)

x_{ij} = j^{th} dose of the i^{th} chemical

β_0 = background response

β_i = slope for the i^{th} chemical

δ = mixture threshold

The data requirements for this model include only the single chemical dose-response curve for each component of the mixture and toxicity data for the same endpoint on the mixture points of interest. Thus, the single chemical and 4-THM experiments were designed specifically to test the hypothesis of dose-additivity for various mixtures of environmental concern (Table II).

This design provides a way to understand the potential interactions among the THMs while avoiding the complications and expense of a larger, more complex study. Figure 1 plots the observed single dose-response curves and the modeled prediction interval for a total mixture dose of 0.872 mmol/kg/d. The filled diamond shape that falls within the 95% prediction interval is the mean experimental SDH level for the mixture point. Because it is within the prediction interval, we cannot conclude that synergism or antagonism are operating for this mixture of environmental concern; such findings support the use of dose-additivity assumptions in risk assessments.

Interaction-Based Hazard Index

The interaction-based hazard index was developed as a way to incorporate information on toxicologic interactions in binary combinations of chemicals into a risk characterization[36, 37]. The approach shown here is a modification of initial work in this area by Mumtaz et al.[38, 39]. This method produces a risk indicator for a group of chemicals that uses known data on chemical pairs and single chemical dose-response information to improve risk indication estimates for the entire mixture. The equation for this method is:

$$HI_{INT} = \sum_{j=1}^{n}(HQ_j * \sum_{j\neq k}^{n} f_{jk}\, M_{jk}^{\varpi jk\, \theta jk}) , \qquad (2)$$

where:

HI_{INT} = hazard index modified by binary interactions data

HQ_j = hazard quotient for chemical j (unitless, e.g., daily intake/RfD)

f_{jk} = toxic hazard of the k^{th} chemical relative to the total hazard from all chemical potentially interacting with chemical j (thus, j cannot equal k)

M_{jk} = interaction magnitude, the influence of a chemical k on the toxicity of chemical j

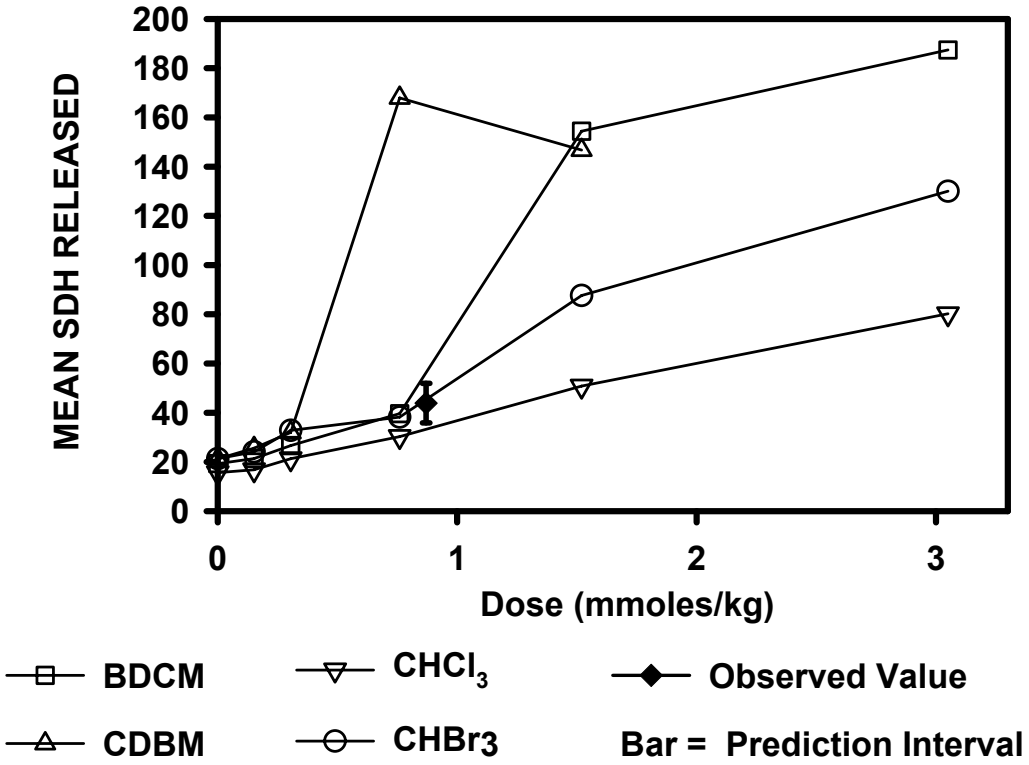
ϖ_{jk} = score for the strength of evidence that chemical k will influence the toxicity of chemical j

θ_{jk} = degree to which chemicals j and k are present in equitoxic amounts

The Hazard Index for dose addition, termed here the additive Hazard Index, is applied only to toxicologically similar chemicals, and is the sum of hazard quotients, which are exposure levels scaled by relative toxic potency[40]. The implementation by the U.S. EPA's Superfund

Figure 1

Additivity of Hepatotoxicity for THMs in Female Mice



Note: Figure 1 plots the observed mean serum SDH (IU/l) for the 4 single THMs and connects them. It compares the prediction interval under dose-additivity with the observed response for the mixture with total dose of 0.872, showing no departure from additivity.

office is to use 1/RfD to scale oral exposures and 1/RfC to scale inhalation exposures[41]. When the additive Hazard Index exceeds 1, there may be a public health concern, so more investigation is warranted. Equation 2 represents a modification of the additive Hazard Index using pairwise interactions, similar to a first-order correction term in linear models or approximation theory. Note that as the interaction magnitude gets smaller (M approaches 1), or the evidence for interaction get weaker (ϖ approaches 0), the right-hand summation approaches 1, and the equation simplifies to the additive Hazard Index.

Each chemical will have its own correction term (the second summation), representing the influence on its toxicity caused by all the other chemicals in the mixture. Each quantity has specific constraints and a formula or rules for its estimation[36]. This second summation is a simple correction formula that is expected to be replaced as better data and information on toxicologic interaction become available. The CD-1 mice data for the single chemicals and the binary mixtures will be used to estimate the interaction-based hazard index for the THMs. The 4-THM mixtures data will then be used to compare the risk indicator with the whole mixture response and further refine the model.

Proportional-Response Addition Method

Proportional-response addition is a component-based risk characterization method that can be applied within a toxicologic context as a substitute for the "usual" methods based on dose addition or response addition[42]. This risk characterization method estimates each component's risk at the total mixture dose and then scales the response proportionately by the amount of the component in the mixture. The basic model for this method for n components is given by equation 3.

$$P_{mix} = \pi_1 P_1(D) + \pi_2 P_2(D) + ... + \pi_n P_n(D) , \quad (3)$$

where:

P_{mix} = Probability of effect for the mixture

$P_i(D)$ = Probability of effect for component i at the total mixture dose D

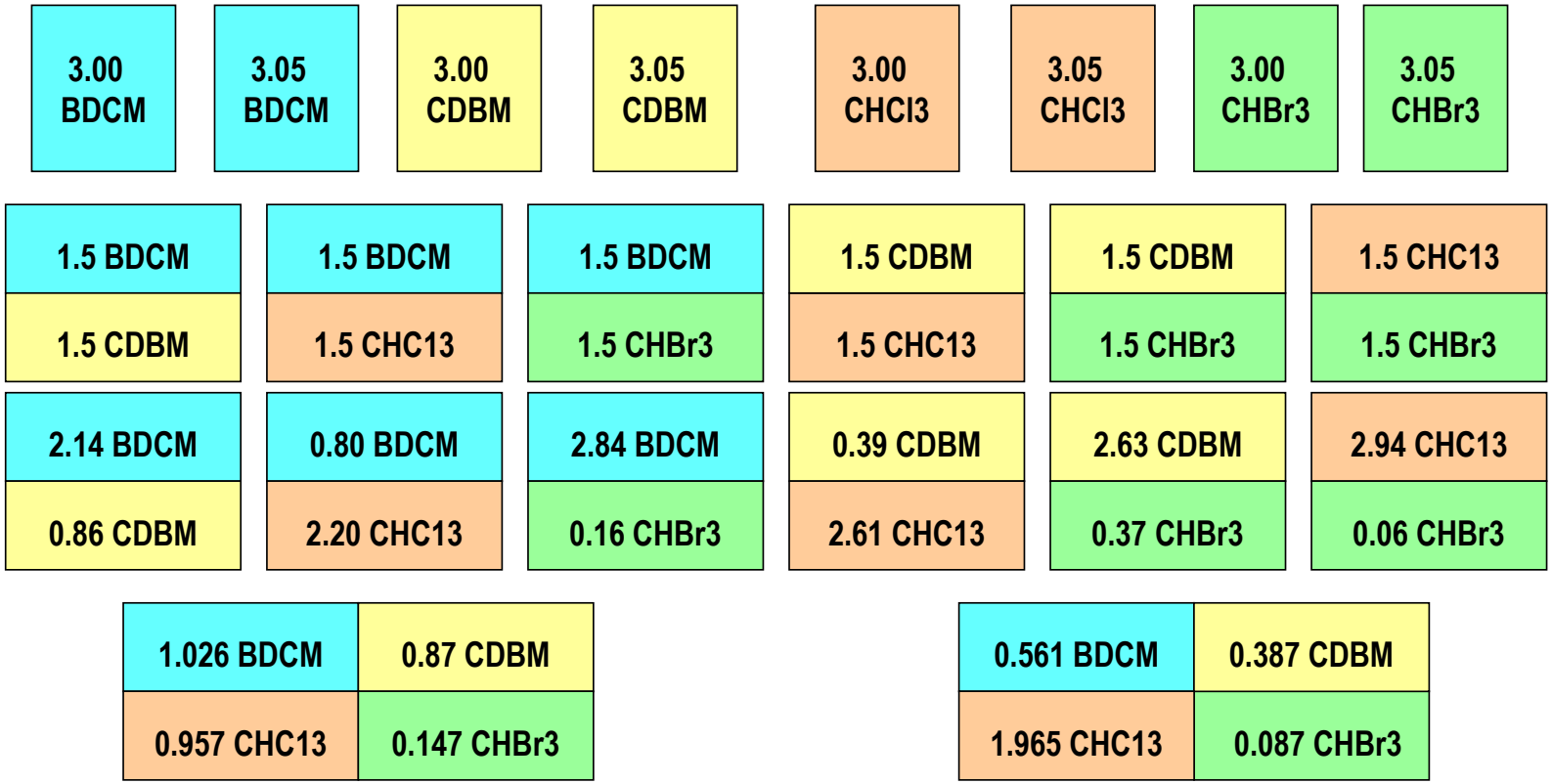
D = Sum of the component doses, not adjusted for toxic potentcy

π_i = Proportion of component i in the mixture

To further develop and refine this method requires a number of data points for which the total mixture dose is held constant, but the proportions of the components of the mixture vary, such as those shown in Tables II and III. Figure 2 illustrates the type of data that will be available for this purpose for a total mixture dose of 3 mmol/kg/d from single chemical,

Figure 2

Single THMs and Mixtures of THMs: Total Dosage ≈3 mmol/kg/day



Note: Each connected group of boxes represents one experimental dose group. Each single THM was tested twice. There are twelve binary combinations and two 4-THM mixtures.

binary and 4-THM experiments. Similar figures could be constructed for the other total dose levels. Proportional-response addition is an interesting risk assessment method because of its potential application to the evaluation of health effects that are thought to be threshold in nature for the individual components of the mixture, but where a lower threshold for the components of chemicals may be expected to exist (e.g., developmental effects that are suggested by the epidemiologic data for drinking water exposures).

Discussion

The multiple-purpose design approach to the DBP mixtures problem uses a multi-disciplinary team of researchers to develop laboratory data for multiple uses. This research effort blends the expertise of laboratory investigators with that of statisticians and risk assessors in order to optimize the utility of the data that are produced.

This paper has detailed the design of research in rodents. Similar experiments are being carried out in the medaka. Since there are potentially hundreds of DBPs of concern in drinking water[5], ways to quickly and inexpensively evaluate and prioritize the toxicity of these compounds are needed. Furthermore, these DBPs must be screened as mixtures for potential interaction effects or additive responses. The medaka is an alternative species that is being studied for use as a possible screening assay for mixtures. Results from medaka embryo assays can yield information on developmental effects, particularly critical circulatory/heart and neurological endpoints, within a 96-hour embryo exposure period. These data may be used to select chemicals for longer-term mammalian studies. Longer-term (one-year) medaka studies focus on reproductive toxicity, cancer and other systemic effects. Thus, the medaka is an important assay for the DBPs in particular because these developmental, reproductive and carcinogenic effects are the endpoints of concern in the epidemiologic literature. In the current research, analogous studies are being conducted in both the CD-1 mouse and medaka assays; cross-species comparisons or corroborations of effects at proportionally equivalent dose levels will be evaluated. Cross species extrapolation evaluation will provide data on this ongoing risk assessment problem and will also aid in the evaluation of the medaka as a screening assay.

The multiple-purpose design approach provides benefits for toxicity evaluation as well as for risk assessment. Testing for departures from additivity examines the validity of the default assumption of dose-additivity in risk assessment. Such studies are efficiently designed to provide interactions information without performing large, expensive bioassays. Experiments that include different "blends" of chemicals and different total mixtures doses will help to demonstrate the utility of proportional-response addition as a method. Single chemical and binary information can be used to estimate the interaction-based hazard index for the THMs, and then the whole mixtures data can be compared to the result to help refine the parameters of that model. The environmentally representative mixing ratios will be useful in the evaluation of health risk for various drinking water treatment scenarios, such as differences between ozonation and chlorination disinfection. Comparisons can be made for the two different species to evaluate cross-species extrapolation issues and to establish the medaka assay as a screening tool for mixtures research. The future of mixtures research will depend on such collaborative efforts that

maximize the use of resources and focus on issues of high relevance to the risk assessment of human health.

Acknowledgements

The authors thank Drs. L.S. Birnbaum, H. Barton, G. Suter, D. DeMarini and Mr. G. Rice for thoughtful review of this manuscript. This research was supported by U.S. EPA cooperative agreements CR-822671, CR-822766 and CR-822761.

References

1. Krasner, S. W., McGuire, M. J., Jacangelo, J. G., Patania, N. L., Reagan, K. M. and Aieta, E. M., The occurrence of disinfection byproducts in U.S. drinking water, *Journal of American Water Works Association*, 81:41-53, 1989.

2. Jacangelo, J. G., Patania, N. L., Reagan, K. M., Aieta, E. M., Krasner, S. W. and McGuire, M. J., Ozonation: Assessing its role in the formation and control of disinfection byproducts, *Journal of American Water Works Association*, 81:74-84, 1989.

3. Lykins Jr., B. W., Koffskey, W. E. and Patterson, K. S., Alternative disinfectants for drinking water treatment, *Journal of Environmental Engineering*, 120(4):745-758, 1994.

4. Miltner, R. J., Rice, E. W. and Stevens, A. A., Pilot-Scale Investigation of the Formation and Control of Disinfection Byproducts. In: 1990 Annual Conference Proceedings. AWWA Annual Conference, Cincinnati, Ohio. 2:1787-1802, 1990.

5. Richardson, S. D., Identification of Drinking Water Disinfection By-Products. In: John Wiley's Encyclopedia of Environmental analysis and Remediation. Robert A. Meyers, Ed. 3:1398-1421, 1998.

6. Bull, R. J., Robinson, M., Meier, J. R. and Stober, J., Use of biological assay systems to assess the relative carcinogenic hazards of disinfection by-products. *Environmental Health Perspectives*, 46:215-227, 1982.

7. Kavlock, R., Chernoff, N., Carver, B. and Kopfler, F., Teratology studies in mice exposed to municipal drinking water concentrates during organogenesis, *Food and Cosmetics Toxicology*. 17:343-347, 1979.

8. Cantor, K. P., Hoover, R. and Hartge, P. et al., Drinking water source and bladder cancer: A case-control study. In: Water Chlorination: Chemistry, Environmental Impact and Health Effects, Vol. 5. Jolley, R.L., R.J. Bull and W.P. Davis et al., Ed. *Lewis Publishers, Inc., Chelsea, MI.* p. 145-152, 1985.

9. McGeehin, M. A. et al., Case-control study of bladder cancer and water disinfection methods in Colorado, *American Journal of Epidemiology*, 138:492-501, 1993.

10. King, W. D. and Marrett, L. D., Case-control study of water source and bladder cancer. *Cancer Causes and Control*, 7:596-604, 1996.

11. Cantor, K. P., Lunch, C. F., Hildesheim, F. et al., Drinking water source and chlorination byproducts. I. Risk of bladder cancer, *Epidemiololgy*, 9:21-28, 1997.

12. Freedman, M., Cantor, K. P., Lee, N. L. et al., Bladder cancer and drinking water: A population-based case-control study in Washington County, Maryland (United States). *Cancer Causes and Control*, 8:738-744, 1997.

13. Bove, F. J., Fulcomer, M. C., Klotz, J. B., Esmart, J., Dufficy, E. M. and Savrin, J. E., Public drinking water contamination and birth outcomes, *American Journal of Epidemiology*, 141(9):850-862, 1995.

14. Kramer, M. D., Lynch, C. F., Isacson, P. and Hanson, J. W., The association of waterborne chloroform with intrauterine growth retardation, *Epidemiology*, 3(5):407-413, 1992.

15. Swan, S. H., Waller, K., Hopkins, B. et al., A prospective study of spontaneous abortion: Relation to amount and source of drinking water consumed in early pregnancy, *Epidemiology*, 9(2):126-133, 1998.

16. Waller, K., Swan, S. H., DeLorenze, G. and Hopkins, B., Trihalomethanes in drinking water and spontaneous abortion, *Epidemiology*, 9(2):134-140, 1998.

17. Gennings, C., Schwartz, P., Carter, W. H. Jr. and Simmons, J. E., Detection of Departures from Additivity in Mixtures of Many Chemicals with a Threshold Model, *Journal of Agricultural Biology Environmental Statistics*, 2:198-211, 1997.

18. Pereira, M. A., Kewa, L. and Kramer, P. M., Promotion by mixtures of dichloroacetic acid and trichloroacetic acid of N-methyl-N-nitrosourea-initiated cancer in the liver of female B6C3F1 mice, *Cancer Letters*, 115:15-23, 1997.

19. Bull, R. J. and Kopfler, F. C., Health Effects of Disinfectants and Disinfection By-Products, *AWWA Research Foundation*, 1991.

20. NTP (National Toxicology Program), Toxicology and carcinogenesis studies of chlorodibromomethane in F344/N rats and B6C3F mice (gavage studies). NTP TR282, 1985.

21. NTP (National Toxicology Program), Toxicology and carcinogenesis studies of bromodichloromethane in F344/N rats and B6C3F mice (gavage studies), *NTP Technical Report*, Ser. No. 321, NIH Publ. No. 87-2537, 1986.

22. NTP (National Toxicology Program), Toxicology and carcinogenesis studies of tribromomethane and bromoform in F344/N rats and B6C3F mice (gavage studies), *NTP*-350, 1989.

23. Smith, M. K., Randall, J. L., Read, E. J. and Stober, J. A., Teratogenic activity of trichloroacetic acid in the rat, *Teratology*, 40:445-451, 1989.

24. Kool, H. J., van Kreijl, C. F. and van Kranen, H. J., The use of XAD-resins for the detection of mutagenic activity in water. II. Studies with drinking water, *Chemosphere*, 10:99-108, 1981.

25. Loper, J. C., Lang, D. R., Schoeny, R. S., Richmond, B. B., Gallagher, P. M. and Smith, C. C., Residue organic mixtures from drinking water show *in vitro* mutagenic and transforming activity, *Journal of the Toxicology and Environmental Health*, 4:919-938, 1978.

26. Nestman, E. R., Otson, R., LeBel, G. L., Williams, D. T., Lee, E. G. H.and Biggs, D. C., Correlation of water quality parameters with mutagenicity of drinking water extracts. In: Water Chlorination: Environmental Impact and Health Effects, Vol. 4. Eds. R.L. Jolley, W.A. Brungs and R.B. Cumming, *Ann Arbor Science Publ., Ann Arbor, MI*, 1982.

27. Weisel, C. P., Kim, H., Haltmeier, P. and Klotz, J. B., Exposure estimates to disinfection by-products of chlorinated debunking water, *Environmental Health Prospectives*, 107:103-110, 1999.

28. Narotsky, M. G., Weller, E. A., Chinchilli, V. M. and Kavlock, R. J., Nonadditive developmental toxicity in mixtures of trichloroethylene, Di(2-ethylhexyl), phthalate and heptachlor in a 5 x 5 x 5 design, *Fundamental and Applied Toxicology*, 27(2):203-16, 1995.

29. Groten, J. P., Schoen, E. D., Van Bladeren, P. J., Frieke Kuper, C., Van Zorge, J. A. and Feron, V., Subacute toxicity of a mixture of nine chemicals in rats: detecting interactive effects with a fractionated two-level factorial design, *Fundamental and Applied Toxicology*, 36:15-29, 1997.

30. Kelly, C. and Rice, J., Monotone smoothing with application to dose-response curves and the assessment of synergism, *Biometrics*, 46:1071-1085, 1990.

31. Haas, C. N. and Stirling, B.A., New quantitative approach for analysis of binary toxic mixtures, *Environmental Toxicology and Chemistry*, 13:149-156, 1994.

32. Gennings, C. and Carter, W. H., Jr., Utilizing concentration-response data from individual components to detect statistically significant departures from additivity in chemical mixtures, *Biometrics*, 51:1264-1277, 1995.

33. Munson, A. E., Sain, L. E., Sanders, V. M. and Kauffmann, B. M. et al., Toxicology of organic drinking water contaminants" trichloromethane, bromodichloromethane, dibromochloromethane and tribromomethane, *Environmental Health Perspectives*, 46: 117-126, 1982.

34. U.S. EPA, Integrated Risk Information System (IRIS). Online. National Center for Environmental Assessment, Cincinnati, OH, 199a.

35. U.S. EPA, Workshop Report: The Risk Assessment of Mixtures of Disinfection By-Products (DBPs) in Drinking Water. NCEA-C-0459, 1998.

36. Hertzberg, R. C., Rice, G. and Teuschler, L. K., Methods for Health Risk Assessment of Combustion Mixtures. In: Hazardous Waste Incineration: Evaluating the Human Health and Environmental Risks. S. Roberts, C. Teaf, and J. Bean eds. *CRC Press* LLC. pp. 105-148, 1999.

37. U.S. EPA, *Guidance for Conducting Health Risk Assessment of Chemical Mixtures*. External Review Draft. NCEA-C-0148, 1999b.

38. Mumtaz, M.M., Durkin, P.R. A weight-of-evidence approach for assessing interactions in chemical mixtures. *Toxicol. Ind. Health*, 8(6): 377-406,1992.

39) Mumtaz, M.M., C.T. DeRosa, J. Groten, U.J. Feron, H. Hansen and P.R. Durkin. Estimation of Toxicity of Chemical Mixtures through Modeling of Chemical Interactions. *Environmental Health Perspectives*, 106(6): 1353-1360, 1998.

40) U.S.EPA, Guidelines for the Health Risk Assessment of Chemical Mixtures. Federal Register.51(185):34014-34025, 1986.

41) U.S. EPA. 1989. Risk Assessment Guidance for Superfund. Vol. 1. Human Health Evaluation Manual (Part A). EPA/540/1-89/002.

42) Chen, J. J., Heflich, R. H. and Bass, B. S., A response-additive model for analyzing mixtures of two chemicals in salmonella reversion assay, *Biomedical Journal*, 31:495-503, 1989.

*Errata**

A Multiple-Purpose Design Approach to the Evaluation of Risks from Mixtures of Disinfection By-Products

Linda K. Teuschler, Chris Gennings, William M. Stiteler, Richard C. Hertzberg, Joan T. Colman, Arunthavarani Thiyagarajah, John C. Lipscomb, William R. Hartley, and Jane Ellen Simmons

Recently we discovered an error in the dosing values of chlorodibromomethane (CDBM), a chemical used in the example mixture described in Teuschler et al. (2000). This note provides the corrected doses for CDBM and indicates the impact this change has on Figures 1 and 2 of the manuscript. Under the Experimental Data section, the doses used in the single chemical experiments for each of the four trihalomethanes (THMs) tested (of which CDBM is one) are given as 0, 0.152, 0.305, 0.76, 1.52 and 3.05 mmol/kg/day. For CDBM, in the single chemical studies only, the correct doses are 0, 0.304, 0.610, 1.52, 3.04 and 6.10 mmol/kg/day.

The dose-response curve for CDBM is shown in Figure 1 of the manuscript, along with the other THMs with the results of the tests for departures from additivity. The change in doses for CDBM results in a flattening of its dose-response curve. A reanalysis of the mixture data points did not reject the hypothesis of dose-additivity, the same result as before finding the error. Thus, these data still support the use of dose-additivity assumptions for the risk assessment of THMs. Finally, it may be noted that the dose for CDBM in Figure 2 of 3.05 mmol/kg/day should now read 3.04 mmol/kg/day.

LK Teuschler, C Gennings, WM Stiteler, RC Hertzberg, JT Colman, A Thiyagarajah, JC Lipscomb, WR Hartley, JE Simmons. 2000. A Multiple-Purpose Design Approach to the Evaluation of Risks from Complex Mixtures of Disinfection By-Products (DBPs*). Drug and Chemical Toxicology*. 23(1):307-321.

* Accepted for Publication in 2000.

ADVANCES IN THE TOXICOLOGICAL ASSESSMENT OF DISINFECTION BY-PRODUCTS IN RODENT AND FISH BIOMEDICAL MODELS*

W. R . Hartley
Department of Environmental Health Sciences, School of Public Health & Tropical Medicine and Environmental Diseases Prevention Research Center, Tulane University Medical Center, Tulane Avenue, New Orleans, LA 70112.

C. Gennings
Department of Biostatistics, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298-0032

L. Teuschler
National Center for Environmental Assessment, US EPA, Cincinnati, OH 45268

A. Thiyagarajah
Department of Environmental Health Sciences, School of Public Health & Tropical Medicine and Environmental Diseases Prevention Research Center, Tulane University Medical Center, Tulane Avenue, New Orleans, LA 70112.

J. E. Simmons
Environmental Toxicology Division, National Health & Environmental Effects Research Laboratory, US EPA, Research Triangle Park, NC 27711

August 1999

**Citation for this work: Hartley, W.R., C. Gennings, L. Teuschler, A. Thiyagarajah, J.E. Simmons. 1999. Advances in the Toxicological Assessment of Disinfection By-products in Rodent and Fish Biomedical Models. AWWA 1999 Annual Conference Proceedings. Chicago, IL.*

Disclaimer

Abstract

The potential health effects of complex mixtures of disinfection by-products in drinking water is a critical public health issue. Typical disinfection by-products (DBP) from treatment of drinking water with halogens include mixtures of chloroform, bromoform, chlorodibromomethane (CDBM), and bromodichloromethane (BDCM). Determination of the toxicological interactions of DBP mixtures is essential in the risk assessment process and selection of drinking water disinfection methods which minimize potential health risks. An array of efficient and short-term toxicological methods for assessing the toxicity of DBP mixtures is presented. The methods presented for assessment of DBPs include the use of the Japanese medaka (*Oryzias latipes*) for assessment of developmental toxicity (neurological, circulatory/heart, and reproductive effects) and cancer, and the rodent for assessment of hepatic and renal toxicity. The proposed methods for the medaka and rodent are demonstrated by toxicological evaluation of bromoform and chloroform. As these toxicological methods are further refined, they may be used to provide toxicological information for risk management decisions to reduce health risks from drinking water with DBP contaminants.

Introduction

With the advent of drinking water disinfection, morbidity and mortality from infectious diseases were greatly reduced and the public health of the nation was greatly improved. The subsequent move to provide safe (disinfected) drinking water nationwide for an increasing population resulted in source water for municipal drinking water from a variety of waterbodies. In particular, surface water sources (rivers, streams and impoundments) resulted in water with multiple carbon sources which upon disinfection with halogens such as chlorine resulted in disinfection by-products (DBPs). Chloroform, bromoform, bromodichloromethane (BDCM), and chlorodibromomethane (CDBM) are DBPs of toxicological concern that occur with common disinfection methods such as chlorination and chloramination. Due to different source water quality, treatment system regimens, and seasonal water quality parameters, the concentrations and relative proportions (ratios) of these four trihalomethanes (THMs) and other DBPs vary significantly.

Low concentrations of DBPs such as chloroform, bromoform, BDCM, and CDBM typically occur in all disinfected drinking water and pose concerns with regard to health risk assessment. In a recent drinking water study, Weisel et al. (1999) found (49 households across New Jersey) mean (+/-sd) concentrations of 31 ± 46 µg $CHCl_3$/L, 0.73 ± 0.90 µg $CHBr_3$/L, 5.7 ± 8.6 µg BDCM /L and 2.0 ± 2.1 µg CDBM/L. Significant indoor air levels of these four DBPs were also found for over 20 minutes post shower use. As complex mixtures, it is unknown if toxicological effects are additive, synergistic, or antagonistic. The effect of complex mixtures of DBPs is a critical public health issue when the ubiquitous human exposure is considered. To demonstrate methods, the authors report some short-term toxicological approaches (USEPA, 1998) to the problem of DBP evaluation using single chemicals and mixtures of chloroform and bromoform.

Methodological Considerations

A first issue in developing toxicological methods to study DBP mixtures was to examine endpoints related to human health. In epidemiological studies in which humans were exposed to DBPs in drinking water, concerns have been raised with regard to the potential for carcinogenicity (Cantor et al., 1985; McGeehin et al., 1993; King and Marrett, 1996; Cantor et al., 1997; Freedman et al., 1997) and reproductive and developmental toxicity (Krasner et al., 1989; Bove et al., 1995; Swan et al., 1998; Waller et al., 1998). Additional concerns were raised by studies in animals of the four THMs (administrated as single chemicals) in which liver and kidney toxicity were prominent (NTP, 1985, 1986, 1989; Smith et al., 1989; Bull and Kopfler, 1991).

A second issue was to determine the type of data needed and the urgency of obtaining information. DBP mixtures are highly variable based on studies of drinking water extracts and exposure studies (Krasner et al., 1989). Additionally there is the lack of scientific methods for evaluation of complex mixtures. The pressure of regulatory decisions and water-treatment technology decisions are also factors that indicate the need for short term and economical

methods to evaluate the critical endpoints of toxicity for DBPs. The authors have tried to address some of these concerns by developing a threshold additivity model that incorporates an efficient experimental design with statistical methodology for prediction of mixture effects under an assumption of dose additivity. The laboratory data used for development of the threshold additivity model was collected in the female CD-1 mouse as this was the only species/strain in which the hepatic toxicity of all four THMs had been assessed by dosing in an aqueous gavage vehicle with a subacute dosing duration (Munson et al., 1982). The authors have also developed toxicological methods for consideration that may relatively efficiently screen mixtures for cancer potential, as well as reproductive and developmental toxicity. A battery of tests is proposed using the Japanese medaka fish (*Oryzias latipes*) as a laboratory model for carcinogenicity and other chronic conditions, and reproductive and developmental effects. Although the female CD-1 mouse was used to collect laboratory data for development of the threshold additivity model, the authors recommend that other rodent strains, the F-344 rat or the B6C3F1 mouse, be used for routine screening of DBPs for hepatic and renal toxicity and for evaluation of the nature of the interaction (additive, synergistic, antagonistic) of mixtures of DBPs. Taken together, the authors are proposing a battery of toxicological tests that can be used to screen mixtures of DBPs for endpoints relevant to human health.

Japanese Medaka

Research fish, particularly medaka, are promising alternative models used in toxicity, carcinogenicity, and developmental toxicity research on environmental chemicals (Smithberg, 1962; Schreiweis and Murray, 1976; Anderson et al., 1977; Llewellyn et al., 1977; Dial, 1978; Solomon and Weis, 1979; Stoss and Haines, 1979; Couch et al., 1981; Couch and Courtney, 1987; Hawkins et al., 1988; Hinton et al., 1988; Hyodo-Taguchi and Egami, 1989; Wisk and Cooper, 1990; Villalobos et al., 1996; Hartley et al., 1998; Lipscomb et al., 1998). The medaka has been used to study cancer and developmental/reproductive toxicity in which the fish are exposed in dynamic (flow-through) systems for one year or more. In considering the medaka for the development of short-term methodology for screening DBPs for cancer and reproductive/developmental toxicity, advantages of limiting exposure to critical periods of embryological development were considered. The advantages of limiting exposure periods to the embryo were: known sensitivity to mammalian carcinogens and reproductive or developmental toxicants; transparency of the chorionic membrane (membrane surrounding the developing embryo) such that developmental effects could be continuously or periodically observed; simplicity of exposure regimens; and the ability to focus on critical periods of development. Focus on a critical developmental stage must be related to important early stages in human development, which occur during the first trimester of pregnancy such as cellular differentiation, as well as development of neurological, circulatory, and hepatic systems. Development of these systems with the exception of time is similar in the medaka and human embryo and are summarized and compared in Table 1. Therefore, the developing medaka embryo (Cameron et al., 1985) would be useful to screen DBPs for effects targeting the developing nervous and circulatory systems. It follows that the exposure period should be as soon as possible after fertilization and be a duration of at least 48 hours. Studies in which the medaka embryo is used

Table 1: Medaka Development at 25°C Compared to Humans (Adapted from Cameron et al., 1985; Kirchen and West, 1976). Times given are approximate post fertilization.

Developmental Characteristics		Medaka (hours)	Humans (days)
(1)	16 cells; single layer of cells, undifferentiated	3	3
(2)	64 cells; two layers of blastomeres surrounded by 16 peripheral cells. Cleavages become asynchronous	4	4
(3)	early high blastula	6	4.5
(4)	late high blastula	9	5
(5)	solid keel of the central nervous system	23	17
(6)	optic vesicle; early eye	29	29
(7)	pericardial cavity; early heart	36	26
(8)	heartbeat; optic lens	46	28
(9)	liver rudiment	128	35

for cancer assessment and other chronic effects should include an exposure period well after the development of the liver rudiment or at least greater than 128 hours (>5 days). The grow-out period should extend for 6 months to 1 year for neoplasm (tumor) development.

The medaka embryo is probably not the best model for evaluating gonadal differentiation and development. In medaka, unlike humans, the newly hatched fry (measuring 4.5-5.4 mm length) have no sexual differences in the histological structure of the somatic elements of the gonads (Satoh and Egami, 1972; Hartley et al., 1998). However, like humans, the medaka is dimorphically stable and intersex fish (hermaphrodites/testis ova) are uncommon in nature. Gonads will differentiate into genetically determined male or female within 2 weeks post hatch. Therefore the exposure time frame to see if DBPs disrupt gonadal development is exposure immediately following hatch and up to 2 weeks post hatch. Disruption of gonadal development can occur by two primary potential mechanisms. The first mechanism is endocrine disruption

where the chemical mimics hormones such as estrogen. The second mechanism could be the DBPs or its metabolites directly acting on gonadal tissues such that undifferentiated gonadal tissue is destroyed or injured by pathological responses such as necrosis or inflammatory reactions. The human reproductive concerns cited above, feminization of fish in proximity of sewage out-falls (Jobling and Sumpter, 1993; Purdom et al., 1994), and initial results from 6-month growouts exposed to chloroform or bromoform showing adverse reproductive effects (U.S. EPA, 1998) support consideration of these biological endpoints in screening for health effects.

Female CD-1 Mice

Hepatic and renal toxicity were selected for assessment as the RfD values for chloroform, CDBM, and bromoform are based on the liver as the critical endpoint for toxicity and the RfD value for BDCM is based on the kidney as the critical endpoint (U.S. EPA, 1999). As mentioned above, the CD-1 mouse was selected based on the primary objective of the rodent research portion of the project: development of data designed to help validate the threshold additivity model for assessment of the additivity assumption for mixtures of chemicals. The secondary objective of the rodent research was to understand the nature of the interactions (additive, synergistic, antagonistic) among the four THMs. Rodents are commonly used to assess hepatic and renal toxicity; an advantage to the use of the rodent model is that endpoints assessed in the rodent can be directly measured in the human. Studies with chloroform and BDCM have shown that clinical chemistry indicators in serum (for hepatic toxicity) and urine (for nephrotoxicity) correlate well with histopathological evidence of hepatic and renal injury (Lilly et al., 1994, 1996, 1999; Thornton-Manning et al., 1994). Serum indicators of hepatotoxicity measured included sorbitol dehydrogenase (SDH), alanine aminotransferase (ALT) and aspartate aminotransferase (AST). Urinary measures of renal toxicity included gamma glutamyl transferase (GGT), alkaline phosphatase (ALP), ALT, AST, and total protein (TPRH).

Chronic studies on rodents typically take one or two years. However, studies have shown that histopathological and enzymatic indicators of toxicity can occur in shorter times such as 14 days and serve as indicators of potential chronic toxicity. Hartley and Ohanian (1988) evaluated uncertainty factors for extrapolation of study doses in short term studies to chronic effects. The ratio of 2-year LOAELS/7 day LOAELS (Lowest Observable Adverse Effect Level) was 0.028 (95th percentile value) representing an uncertainty factor of 35. Although 14-day repeated exposure studies in rodents will probably never be used to set reference doses (safe lifetime doses for humans), they are extremely useful to identify the potential long term health effects of DBPs.

Results - Medaka

Developmental Toxicity

A culture of Japanese medaka was maintained under optimal conditions (Kirchen and West, 1976; Hartley et al., 1995) to obtain stage-specific embryos. Embryos were obtained by

removal of batches of eggs from females and removal of chorionic filaments by mechanical abrasion. Embryos were typically in the 16-cell stage. Batches of eggs were pooled and checked for fertilization. Unfertilized eggs were identified by a cloudy chorion and discarded. By random selection, each embryo was put into an individual glass vial with 5 mL of embryo-rearing solution, sealed and put into a water bath maintained at 25°C. The embryo rearing solution was prepared according to Rugh (1962). Collection, screening and handling of the embryos results in exposure beginning in the early high blastula stage. Embryos were examined daily by light microscope with video monitor for lethality, and embryological development with focus on the circulatory and nervous systems development (Table 2). Heartbeat for each embryo was measured for one minute by observation on the microscope video monitor and using an audible (alarm) timer.

For developmental studies, embryos were exposed for a period of 4 days to the DBP. There was a complete volume exchange every day to ensure consistent exposure and reduce the effects of volatility. Embryos (10/exposure concentration) were exposed to single DBPs to determine the 96-hour LC_{50} (concentration lethal to 50% of the embryos). For detailed developmental studies, time and personnel constraints limit observation to 80 embryos per day. Increased sample size could be obtained by conducting experiments in blocks. Determination of the LC_{50} was based on observed lethality and neurological and circulatory effects so severe that lethality was expected. Exposure concentrations for subsequent developmental studies were selected based on the results of the lethality studies. Exposure solutions were made by adding the DBP (single chemical or mixture) under study to embryo-rearing solution. Stock solutions of the DBPs were prepared with dimethyl sulfoxide (DMSO) and aliquots pipetted into each vial to achieve the desired concentration. Exposure concentrations were expressed as nominal concentrations.

Single DBPs- Chloroform and Bromoform

The 96-hour LC_{50} for chloroform (Table 3) was determined in two blocks with 10 embryos per concentration exposed to 0, 25, 50, 75, 100 or 125 mg/L chloroform in embryo-rearing solution. The LC_{50} was 59 mg/L (95%CL = 50-72 mg/L) based on the combined incidence of mortality and severe neurological/circulatory effects expected to result in death. There was a dose-related increase in combined mortality and neurological and circulatory defects with incidence reaching 100% at concentrations >100 mg/L. Mean heartbeat rate was measured in surviving embryos at 72 and 96 hours. At 72 hours, mean heartbeat rate decreased with
increasing concentration in all dose groups compared to the controls. The heartbeat rate at 72 hours was significantly different in the two high concentration groups. At 96 hours, the heartbeat rate was statistically significantly reduced at the three middle concentrations (50, 75, 100 mg/L), but not at either the lowest (25 mg/L) or highest concentration (125 mg/L). A possible explanation of the high concentration results is that the three embryos alive at 96 hours were hardy survivors.

Table 2

Sequential Abnormal Developmental Endpoints Evaluated

Time	Abnormalities
0 hrs	
24 hrs	Blastoderm becomes an abnormal mass / embryo death Failure of the development of solid keel of the CNS Abnormal solid keel of the CNS Abnormal or no yolk plug formation
48 hrs	Failure to develop or abnormal development of the heart (tubular heart) Absence of heartbeat Lack of blood islands in the yolk sac Failure of the forebrain and optic lens to develop Failure of the head to lift from the yolk sac
72 hrs	Absence of pigmented blood Lack of development and/or orientation of the heart chambers Heartbeat measurement Absence of adequate cardinal vein development Absence of retinal pigments / abnormal eye development Absence of tail and other embryo movement Absence of or abnormal otoliths
96 hrs	Absence of sinuous, cardinal, and vitelline veins Absence of whole animal movement by tail lashing Absence of generalized pink blood Lack of /abnormal heart chambers formation

Table 3

Mortality, Developmental Effects, and Heartbeat Rate in Chloroform-Treated Embryos

Concentration (ppm)	Total N	Mortality	Combined Mortality and Neurological/Circulatory Defects[+]	Heartbeat (beats/min ±SD)	
				At 72 hours	At 96 hours
0	20	0%	0%	80.6±8.7	96.3±5.8
25	20	0%	5%	78.7±11.7	96.8±5.3
50	10	0%	40%	72.5±9.5	79.2±15.2**
75	10	0%	40%	70.3±15.1	84.6±10.6*
100	10	40%	100%	51.7±6.1**	60.3±9.5**
125	10	70%	100%	46.67±9.2**	91.7±11.9

*p<0.05, Tukey A test
**p<0.01, Tukey A test
[+] Severe defects and expected lethality

The 96 hour LC_{50} for bromoform (Table 4) was determined in two blocks with 10 embryos per concentration exposed to 0, 10, 25, 40, 50 or 60 mg/L bromoform in embryo-rearing solution. The LC_{50} (based on combined mortality and severe neurological/circulatory effects) was 15.7 mg/L (95%CL = 6-21 mg/L). There was an increase in combined mortality and neurological and circulatory defects in all exposed groups. At 10 and 25 mg/L, the incidence for these effects was 50-60%, whereas the incidence reached 100% at concentrations > 40 mg/L. Heartbeat rate was measured in surviving embryos at 72 and 96 hours. These results were not consistent in that they did not follow an expected dose-response trend. At 72 hours, mean heartbeat rate was statistically significantly increased at 10 mg/L and significantly decreased at 50 mg/L. The remaining concentration groups (25, 40, and 60 mg/L) had heartbeat rates comparable to the controls. At 96 hours, heartbeat rate was significantly increased at 10 and 25 mg/L and significantly decreased at 60 mg/L.

Table 4

Mortality, Developmental Effects and Heartbeat Rate in Bromoform-Treated Embryos

Concentration (ppm)	Total N	Mortality	Combined Mortality and Neurological/Circulatory Defects[+]	Heartbeat (beats/min ±SD)	
				At 72 hours	At 96 hours
0	20	0%	5%	82.1±10.4	93.1±10.3
10	10	20%	50%	96.8±7.65**	111.4±8.0**
25	20	30%	60%	87.9±10.73	106.8±8.4**
40	10	80%	100%	77.0	100
50	10	70%	100%	65.3±9.5**	101±8.2
60	10	80%	100%	75	64±26.9**

*p<0.05, Tukey A test
**p<0.01, Tukey A test
[+] Severe defects and expected lethality

Binary Mixtures of Chloroform and Bromoform

A mixture experiment with $CHCl_3$:$CHBr_3$ combination points each of a 1:1 ratio was conducted to detect departures from additive toxicological response. Groups of 10 embryos were exposed for 96 hours to 1:1 binary mixtures of $CHCl_3$:$CHBr_3$ at four concentrations of 10:10, 15:15, 20:20 or 25/25 mg/L. Controls consisted of DMSO vehicle controls and positive controls of 75 mg $CHCl_3$/L and 25 mg $CHBr_3$/L. The four combination points (Table 5) as well as the positive controls were selected from the single chemical dose response curves. The statistical models employed for predicting departures from additivity are presented in U.S. EPA (1998) and Gennings et al. (1999; see Section 5 of this report). The type of neurological and developmental defects observed in the mixtures groups were similar to those observed in the single chemical groups. The results of the positive controls in these mixtures experiments were not different from those expected based on the single chemical dose-response curves. Departure from additivity (antagonism) was observed in the highest mixture dose group (25:25 mg/L). There was no departure from additivity for the three lower concentration mixtures groups.

Table 5

Observed and Model Predicted Responses Under Additivity Using Logistic Model for Chloroform:Bromoform Mixtures

	Observed Death/Defect* (%)	Predicted Death/Defect Under Additivity	95% Prediction Interval
Controls			
Control	0.2	0.024	[0.001, 0.370]
$CHCL_3$: 75 ppm	0.4	0.607	[0.248, 0.878]
$CHBr_3$: 25 ppm	0.6	0.669	[0.350, 0.884]
$CHCL_3$: $CHBr_3$ Mixtures			
10:10 ppm	0.4	0.198	[0.042, 0.583]
15:15 ppm	0.3	0.442	[0.161, 0.764]
20:20 ppm	0.4	0.716	[0.341, 0.925]
25:25 ppm	0.2	0.890	[0.461, 0.987]

*N=10/group

Reproductive Toxicity

A novel short-term methodology is proposed in which newly hatched medaka fry are exposed for 48 hours to DBPs to determine if they disrupt normal gonadal development. The presence of testis-ova (hermaphroditism) and sex ratios with a disproportionate number of females would indicate that the DBPs are mimicking estrogen (i.e., endocrine disruptors). Abnormal development of testis with histopathological injury such as necrosis, abnormal sperm or ova, and other effects would indicate that the DBPs are acting directly on the gonads rather than via the endocrine system. In the case of endocrine disruption, the method requires a positive control for human estrogen (17-beta estradiol). Evaluation of the sensitivity of the newly hatched medaka fry to short-term (48-hour) exposure to 17-beta estradiol has been completed. Recently

Hartley et al. (1998) exposed newly hatched medaka fry (30 fry per dose group) with undifferentiated gonads to 4.0, 29.4, and 115.6 µg/L 17-beta estradiol (acetone carrier) for 48 hours in a water bath at 25°C. The fry were grown-out in spring water for 2 weeks, killed, and processed for histopathological evaluation (serial sections). Fry exposed to 17 beta-estradiol developed primarily into females or had testis-ova. With the development of methodology and the establishment of 17-beta estradiol as a positive control, we plan to evaluate the reproductive toxicity and endocrine disruption of individual DBPs and DBP mixtures. In the application of this methodology to DBPs, reproductive toxicity will be functionally defined as: failure of medaka fry to survive; predominantly female and/or hermaphroditic offspring; histologically abnormal testis or ovaries; ectopic (abnormal location) gonads, and gross reproductive abnormalities that would prevent successful breeding. Initially, the same concentrations of chloroform, bromoform, and binary mixtures of both will be used in the further evaluation of this methodology for reproductive toxicity.

Chronic Toxicity

Embryos were collected from the medaka colony in the same manner and stage of development as described for developmental studies. Exposure was continued for 10 days, after which the embryos were rinsed in embryo-rearing solution and transferred for hatching (embryo-rearing solution) and grow-out chambers (spring water). At 6 months and 1 year post hatch, fish were killed with an overdose of Tricane methane sulfonate (MS222), examined for gross morphology, weighed, measured, and fixed in 10% buffered formalin. Samples were decalcified, sectioned using a modification of the method of Wolfe (1994), and stained with hematoxylin and eosin for examination by light microscopy. A complete histopathological examination was performed on all surviving fish. Slides were read blind (i.e., exposure status is unknown to the pathologist), and all lesions, including preneoplastic and neoplastic lesions, were identified.

Results on 20 of 60 embryos per concentration exposed to 0, 25, 50, and 100 mg/L chloroform were reported (U.S. EPA, 1998). At six months there were no neoplasms and probable preneoplastic lesions except 2/20 medaka in the 50mg/L-dose group with cysts and one case of clear cell foci at 100 mg/L. In limited sampling of the chloroform-exposed fish there were chronic cardiac lesions (dilated atrium and tubular heart) and reproductive abnormalities (in some cases, immature ovaries and testes). These chronic conditions were not evident in the control fish. As reported in U.S. EPA (1998) evaluation of twenty of sixty embryos exposed to 0, 10, and 50 mg/L bromoform, at 6 months liver lesions included cysts, spongiosis hepatis, altered foci and clear cell foci. Similar to chloroform in limited sampling, cardiac lesions and reproductive abnormalities were observed.

Results – CD-1 Female Mouse

Liver and kidney toxicity were the basis of the reference dose (safe lifetime dose) for the DBPs under study (U.S. EPA, 1998). The female CD-1 mouse was selected for validation of the threshold additivity model. Expression of toxicity at the histopathological level was typically centrilobular necrosis of the liver. To assess the toxicity of the single DBPs, one study for each

DBP was conducted. The general methods developed and the results for chloroform and bromoform are presented. Experimental doses were selected in accordance with the needs for development of the threshold additivity model for mixture assessment. Doses were selected that were likely to be less than or greater than the anticipated threshold.

For each single DBP, female CD-1 mice (N=8-20), 65-70 days old, were exposed by oral gavage for 14 days to each DBP. Animals were housed in metabolism cages (2/cage) and urine was collected for 24 hr prior to the first day of dosing and on days 0, 1, 7 and 14. Each experiment was conducted in two blocks with part of the animals from each dose group in each block. Each DBP was administered at the following dosages in an aqueous vehicle (10% Alkamulus EL-620Ò) at a constant volume of 10ml/kg: 0, 0.152, 0.305, 0.76,1.52 or 3.05 mmol/kg/day. Dosing solutions were made daily and animals were dosed between 8:00 am and noon, and were killed during the same time interval on Day 15. At termination, serum and urine endpoints and body and organ (liver and kidney) weights were evaluated. Urine was analyzed for indicators of nephrotoxicity: GGT, ALP, ALT, AST, and TPRH. Serum was analyzed for indicators of hepatic toxicity: SDH, ALT, and AST. Tissues (liver and kidney) were preserved in 10% phosphate-buffered formalin for histopathological evaluation. Two of six slides of liver and kidney were stained with hematoxylin and eosin and examined blind (without knowledge of the dose group) by brightfield microscopy. Centrilobular necrosis was characterized as none (no necrotic hepatocytes seen), mild (less than 10% of the hepatocytes involved), mild-moderate (less than 50% of the hepatocytes involved), moderate (approximately 50% of the hepatocytes involved), moderate-severe (more than 50% of the hepatocytes involved), and severe (more than 75% of the hepatocytes involved).

For chloroform, relative liver weight and AST were increased at 1.52 mmol/kg/d. Significant increases in serum SDH and ALT occurred at >0.76 mmol/kg/d. There were dose-related increases in the incidence and severity of centrilobular necrosis. Mild to moderate necrosis occurred at 0.305 mmol/kg/d which progressed in severity with dose from moderate to severe necrosis with severe Zone 2 vacuolation at the highest dose. The No Observed Adverse Effects Level (NOAEL) for centrilobular necrosis was 0.152 mmol/kg/d. Chloroform had no detectable effect on either body weight or relative kidney weight. TPRH was decreased at three lower dosages (0.152, 0.305, 0.76 mmol/kg/d) on Day 7 and two upper dosages, 1.52 and 3.05 mmol/kg/d on Day 14.

For bromoform, relative liver weight was increased at 0.305 mmol/kg/d. There were significant increases in serum SDH and ALT at 0.76 mmol/kg/d, and AST at 1.52 mmol/kg/day. There were dose-related increases in the incidences and severity of centrilobular necrosis. At 0.305 mmol/kg/d, centrilobular necrosis was predominantly mild to moderate, whereas at the two highest doses, moderate to severe cases occurred accompanied by Zone 2 vacuolation. The NOAEL for centrilobular necrosis was 0.152 mmol/kg/d. Bromoform had no detectable effect on body weight. Relative kidney weight was significantly increased at the highest dose. TPRH was decreased on Days 7 and 14 at 3.05 mmol/kg/d.

Discussion

The authors have proposed an array of efficient and short-term toxicological methods using the medaka as a biomedical model to determine the potential human health effects of DBPs. The CD-1 mouse was used, for reasons described above, for validation of the threshold additivity model. However, the authors highly recommend that other more genetically stable rodent strains be used to screen THMs and DBPs for hepatic and renal toxicity and for determination of the nature of THM and DBP interactions with the threshold additivity model. The methods described here cover the range of critical human health endpoints of concern for DBPs including developmental toxicity {medaka} (neurological, circulatory/heart and reproductive effects), chronic conditions {medaka} (cancer, heart, and reproductive effects), and hepatic and renal toxicity {mouse}. The use of this battery of toxicological methods may provide a comprehensive approach for the evaluation of DBPs on many levels. Toxicological methods and procedures presented may meet specific needs in the risk assessment of DBPs.

The identification of the critical toxicological effect (organ system) of single chemical DBPs and mixtures of DBPs is essential and this battery of procedures addresses major toxicological endpoints of concern. The dose-response definition of single chemical DBPs can be used to design mixtures studies to determine if the interactions are additive, synergistic, or antagonistic. This type of basic information is useful in the design of long term studies of prototype mixtures of DBPs. Ranking the toxicity of mixtures of DBPs produced by different disinfection methods, source waters and water treatment facilities will be extremely useful in optimizing treatment selection to reduce the levels of DBPs in general and reducing the fraction of the most toxic DBPs in drinking water. Providing baseline toxicity data on individual DBPs and mixtures of DBPs will ensure that resources for longer term and more expensive chronic studies are well planned and efficiently used to study the most toxic individual DBPs and/or DBP mixtures.

We recommend that short term toxicological methods be used to test DBP mixtures produced by different treatment trains and source waters. The authors have started to employ these procedures to evaluate some prototype mixtures with the mixing ratios, of DBPs based on the average seasonal proportions of 35 water treatment facilities (Krasner et al., 1989). Other studies are in progress evaluating the interaction of binary mixtures of DBPs and comparing the toxicity of mixtures of DBPs produced by chlorination and ozonation (with chloramination). As these methods are further refined, they may be used to make reasonable management decisions to reduce health risk from drinking water with DBP contaminants in a relatively swift and economical manner.

Acknowledgements

The authors thank Drs. L.S. Birnbaum, B. Bruce, L. Hall, J. McKinney, P. Schmieder, R. Johnson, and S. Broderius for thoughtful review of this manuscript. This research was supported by U.S. EPA cooperative agreement numbers CR-822671 and CR 822766-01-4. Portions of the

data reported here are also reported in another paper in this report (Gennings et al., 1999, Section 5) to illustrate the development of novel statistical approaches for assessment of additivity.

References

Anderson, J.W., Dixit, D.B., Ward, G.S., and Foster, R.S. 1977. Effects of petroleum hydrocarbons on the rate of heart beat and hatching success of estuarine fish embryos. In: F.J. Vernberg, A. Calabrese, F.P., Thurberg, and W.B. Vernberg, eds., Physiological Responses of Marine Biota to Pollution, Academic Press, New York. p. 241-258.

Bove, F.J., Fulcomer, M.C., Klotz, J.B., Esmart, J., Dufficy, E.M., and Savrin, J.E. 1995. Public drinking water contamination and birth outcomes. *American J. of Epidemiology* 141:850-862.

Bull, R.J., and Kopfler, F.C. 1991. Health effects of Disinfectants and Disinfection By-Products. AWWA Research Foundation.

Cameron, I.L., Lawrence, W.C., and Lum, J.B. 1985. Medaka eggs as a model system for screening potential teratogens. In: Prevention of Physical and Mental Congenital Defects, Part C: Basic and Medical Sciences, Education, and Future Strategies, Alan R. Liss, Inc. p. 239-243.,

Cantor, K.P., Hoover, R., Hartge, P., et al. 1985. Drinking water source and bladder cancer: a case-control study. In: Jolley, R.L., Bull, R.J., Davis, W.P., et al., (Eds.), Water Chlorination: Chemistry, Environmental Impact and Health Effects, vol 5. Lewis Publishers, Inc., Chelsea, MI.. p. 145-152.

Cantor, K.P., Lunch, C.F., Hildesheim, M., Dosemeci, M., Lubin, J., Alavanja, M., and Craun, G.F. 1997. Drinking water source and chlorination byproducts. I. Risk of bladder cancer. *Epidemiology* 9:21-28.

Couch, J.A., and Courtney, L.A. 1987. N-Nitrosodiethylamine-induced hepatocarcinogenesis in estuarine sheepshead minnow (*Cyprinodon variegatus*): neoplasms and related lesions compared with mammalian lesions. *JNCI* 79:297-321.

Couch, J.A., Courtney, L.A., and Foss, S.S. 1981. Laboratory evaluation of marine fishes as carcinogen assay subjects. TN: Dawe, C.J., Harshbarger, J.C., Kondo, S., et al. Eds. Phyletic Approaches to Cancer. Tokyo, Japan Sci. Soc. Press 125-139.

Dial, N.A. 1978. Methylmercury: some effects on embryogenesis in the Japanese medaka, *Oryzias latipes*. *Teratology* 17: 83-92.

Freedman, M., Cantor, K.P., Lee, N.L., Chen, L.S., Lie, H.H., Ruhl, C.E., and Wang, S.S. 1997. Bladder cancer and drinking water: A population-based case-control study in Washington Country, Maryland (United States). *Cancer Causes and Control* 8:738-744.

Gennings, C., Teuschler, L., Hartley, W.R., Thiyagarajah, A., and Simmons, J.E. 1999. Novel statistical methods for risk assessment of disinfection by-product mixtures. American Water Works Association Annual Conference Proceedings. Chicago, IL.

Hartley, W.R., and Ohanian, E.V. 1988. The use of short-term toxicity data for prediction of long-term health effects, In: Trace Substances in Environmental Health XXII, D.D. Hemphill, ed., University of Missouri Press, Columbia. p. 3-12.

Hartley W.R., Thiyagarajah A., and Mizell M. 1995. Ectopic hermaphroditic gonad in the cranial cavity of Japanese medaka. *Journal of Aquatic Animal Health* 7:172-177.

Hartley, W.R., Thiyagarajah, A., Anderson, M.B., Broxson, M.B., and Major, S.E. 1998. Gonadal development in Japanese medaka (*Oryzias latipes*) exposed to 17 b-estradiol. *Marine Environmental Research* 46:145-148.

Hawkins, W.E., Overstreet, R.M., and Walker, W.W. 1988. Carcinogenicity tests with small fish species, *Aquatic Toxicology* 11: 113-128.

Hinton, D.E., Couch, J.C., Teh, S., and Courtney, L., 1988. Cytological changes during progression of neoplasia in selected fish species. *Aquatic Toxicology* 11:77-112.

Hyodo-Taguchi, Y., and Egami, N. 1989. Use of small fish in biomedical research with special reference to inbred strains of medaka. In: A.D.Woodhead (ed.), Nonmammalian Animal Models for Biomedical Research. CRC Press, Boca Raton, FL. p. 186-204.

Jobling, S., and Sumpter, J.P. 1993. Detergent components in sewage effluent are weakly estrogenic to fish: An *in vitro* study using rainbow trout (*Oncorhynchus mykiss*) hepatocytes. *Aquatic Toxicology* 27: 361-372.

King, W.D., and Marrett, L.D. 1996. Case-control study of water source and bladder cancer. *Cancer Causes and Control* 7:596-604.

Kirchen, R. and West, W., 1976. The Japanese Medaka Its Care and Development, Carolina Biological Supply Company, Burlington, NC, p. 36.

Krasner, S.W., McGuire, M.J., Jacangelo, J.G., Patania, N.L., Reagan, K.M., and Aieta, E.M. 1989. The occurrence of disinfection by-products in U.S. drinking water. *JAWWA* 81:41-53.

Lilly, P.D., Simmons, J.E., and Pegram, R.A. 1994. Dose-dependent vehicle differences in the acute toxicity of bromodichloromethane. *Fundamental and Applied Toxicology* 23: 132-140.

Lilly, P.D., Simmons, J.E., and Pegram, R.A. 1996. Effect of subchronic corn oil gavage on the acute toxicity of orally administered bromodichloromethane. *Toxicology Letters* 87: 93-102.

Lilly, P.D., Sey, Y.M., Seely, J.C., et al. 1999. The influence of gavage vehicle and concurrent exposure to trichloroethylene and chloroform hepatic and renal toxicity. (Submitted).

Lipscomb, J.C., Confer, P.D., Miller, M.R., Stamm, S.C., Snawder, J.E., and Stelvio, M.B. 1998. Metabolism of trichloroethylene and chloral hydrate by the Japanese medaka (*Oryzias latipes*) *in vitro*. *Environmental Toxicology and Chemistry* 17: 325-332.

Llewellyn, G.C., Stephenson, G.A., and Hoffman, J.W. 1977. Aflatoxin B1 induced toxicity and teratogenicity in Japanese medaka eggs (*Oryzias latipes*). *Toxicon* 15: 582-587.

McGeehin, M.A. et al. 1993. Case-control study of bladder cancer and water disinfection methods in Colorado. M. J. *Epidemiology* 138:492-501.

Munson, A.E., Sain, L.E., Sanders, V.M., Kauffmann, B.M., White Jr., K.L., Page, D.G., Barnes, D.W., and Borzelleca, J.F. 1982. Toxicology of organic drinking water contaminants: Trichloromethane, bromodichloromethane, dibromochloromethane, and tribromomethane. *Environmental Health Perspectives* 46: 117-126.

NTP. (National Toxicology Program) 1985. Toxicology and carcinogenesis studies of chlorodibromomethane in F433/N rats and B6C3F1 mice (gavage studies). NTP TR282.

NTP. (National Toxicology Program) 1986. Toxicology and carcinogenesis studies of bromodichloromethane in F344/N rats and B6C3F mice (gavage studies). NTP Technical Reports, Ser. No. 321, NIH Publ. No. 87-2537.

NTP. (National Toxicology Program) 1989. Toxicology and carcinogenesis studies of tribromomethane and bromoform in F344/N rats and B6C3F mice (gavage studies). NTP-350.

Purdom, C.E., Hardiman, P.A., Bye, V.J., Eno, N.C., Tyler, C.R., and Sumpter, J.P. 1994. Estrogenic effects of effluents from sewage treatment works. *Chemical Ecology* 8: 275-285.

Rugh, R. 1962. Experimental Embryology Techniques and Procedures, 3rd ed., Burgess Publishing Co., Minneapolis.

Satoh, N., and Egami, N. 1972. Sex differentiation of germ cells in the teleost, *Oryzias latipes*, during normal embryonic development. *Embryology and Experimental Morph* 28: 385-395.

Schreiweis, D.O., and Murray, G.J. 1976. Cardiovascular malformations *in Oryzias latipes* embryos treated with 2,4,5-trichlorophenoxyacteic acid (2,4,5-T). *Teratology* 14: 287-290.

Smith, M.K., Randall, J.L., Read, E.J., and Stober, J.A. 1989. Teratogenic activity of trichloracetic acid in the rat. *Teratology* 40:445-451.

Smithberg, M. 1962. Teratogenic effects of tolbutamide on the early development of the fish, *Oryzias latipes*. *American Journal of Anatomy* 111: 205-213.

Solomon, H.M., and Weis, J.S. 1979. Abnormal circulatory development in medaka caused by the insecticides carbaryl, malathion and parathion. *Teratology* 19: 51-62.

Stoss, F.W., and Haines, T.A. 1979. The effects of toluene on embryos and fry of the Japanese medaka *Oryzias latipes* with a proposal for rapid determination of maximum acceptable toxicant concentration. *Environmental Pollution* 13: 139-148.

Swan, S.H., Waller, K., Hopkins, B., Windham, G., Fenster, L., Schaefer, C., and Neutra, R.R. 1998. A prospective study of spontaneous abortion: relation to amount and source of drinking water consumed in early pregnancy. *Epidemiology* 9:126-133.

Thornton-Manning, J.R., Seely, J.C., and Pegram, R.A. 1994. Toxicity of bromodichloromethane in female rats and mice after repeated oral dosing. *Toxicology* 94: 3-18.

U.S. EPA. 1998. Workshop Report: The Risk Assessment of Mixtures of Disinfection By-Products (DBPs) in Drinking Water. NCEA-C-0459.

U.S. EPA. 1999. Integrated Risk Information System (IRIS). Online. National Center for Environmental Assessment, Washington, DC.

Villalobos, S.A., Anderson, M.J., Denison, M.S., Hinton, D.E., Tullis, K.T., Kennedy, I.M., Jones, A.D., Chang, D.P.Y., Yang, G., and Kelly, P. 1996. Dioxinlike properties of a trichlorethylene combustion-generated aerosol. *Environmental Health Perspectives* 104: 734-743.

Waller, K., Swan, S.H., DeLorenze, G., and Hopkins, B. 1998. Trihalomethanes in drinking water and spontaneous abortion. *Epidemiology* 9:134-140.

Weisel, C.P., Kim, H., Haltmeier, P., and Klotz, J.B. 1999. Exposure estimates to disinfection by-products of chlorinated drinking water. *Environmental Health Perspectives* 107: 103-110.

Wisk, J.D., and Cooper, K.R. 1990. The stage specific toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in embryos of the Japanese medaka (*Oryzias latipes*). *Environmental Toxicology and Chemistry* 9: 1159-1169.

Wolfe, M.J. 1994. Pathology Assessment Issues. In: Compendium of the FY 1990 & FY 1992 Research Reviews for the Research Methods Branch, U.S.Army Biomedical Research & Development Laboratory, Fort Detrick, Frederick, MD. H.S. Gardner, Jr., Compiler, Technical Report 9404, p. 83-87.

NOVEL STATISTICAL METHODS FOR RISK ASSESSMENT OF DISINFECTION BY-PRODUCT MIXTURES*†

C. Gennings
Dept. of Biostatistics, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298-0032

L. Teuschler
National Center for Environmental Assessment, U.S. EPA, Cincinnati, OH

W. Hartley and R. Thiyagarajah
School of Public Health and Tropical Medicine, Tulane Medical Center, Tulane University, New Orleans, LA

J.E. Simmons
Experimental Toxicology Division, National Health and Environmental Effects Research Laboratory, U.S. EPA, Research Triangle Park, NC

August 1999

**Citation for this work: Gennings, C., L. Teuschler, W. Hartley, A. Thiyagarajah, J.E. Simmons. 1999. Novel Statistical Methods for Risk Assessment of Water Disinfection By-products. AWWA 1999 Annual Conference Proceedings. Chicago, IL.*

†Also see Errata (2000) of Gennings et al. (1997) in Appendix B

Disclaimer

Abstract

Investigation of the assumption of additivity has been hampered by the lack of adequate and appropriate experimental designs and statistical methods. This is particularly true for mixtures of more than two chemicals. Here, we describe the threshold additivity model (Gennings et al., 1997), a flexible experimental design and statistical methodology that is applicable to mixtures of large number of chemicals. Advantages of this design and analytic approach are that it allows investigators to focus on particular mixture combination points of interest, decreasing the number of animals required as well as the time and cost of performing the experiments. This approach is illustrated by two examples applicable to drinking water disinfection by-products formed during either chlorination of water or ozonation of water followed by post-treatment with either chlorine or chloramine. In the first example, the additivity assumption is examined, with hepatotoxicity in female CD-1 mice as the endpoint, for a mixture of the four trihalomethanes, chloroform ($CHCl_3$), bromoform ($CHBr_3$), bromodichloromethane and chlorodibromomethane, formed during disinfection of water. The mixture tested was based on the average seasonal proportions of these four chemicals at 35 water treatment facilities (Krasner et al., 1989). For the particular mixture and dosage tested, the experimental sample mean was within the 95% prediction interval from the threshold additivity model, providing evidence that dose additivity is a reasonable assumption for risk assessment. In the second example, the additivity assumption for developmental toxicity is examined for binary mixtures of $CHCl_3$ and $CHBr_3$ in medaka fish. With the statistical power afforded by the present experiment, antagonism was detected at the highest mixture dose tested (25 ppm $CHCl_3$: 25 ppm $CHBr_3$) and departure from additivity was not detected at the three lower-dose mixture groups. In summary, a threshold additivity model for efficiently detecting departure from additivity is described and illustrated with two biological examples.

Introduction

It is often of interest to study the effects of exposure to multiple substances/chemicals. Of primary importance in such studies is the determination and characterization of interactions among the components in a mixture. For example, in the study of disinfectant by-products (DBPs) found in drinking water, it is of interest to determine if the DBPs interact in such a way as to increase the toxic effect above what one would expect to observe from any single chemical. If the DBPs do not interact, it is said that they may either be functionally independent or may act in a dose-additive fashion. One definition of additivity is given by Berenbaum (e.g., 1985) and is based on the classical isobolograms for the combination of two chemicals (e.g., Loewe and Muischnek, 1926; Loewe, 1953). That is, in a combination of c chemicals, let X_i represent the concentration/dose of the i^{th} component alone that yields a fixed response, y_0, and let x_i represent the concentration/dose of the i^{th} component in combination with the c agents that yields the same response. According to this definition of additivity if the substances combine with zero interaction, then

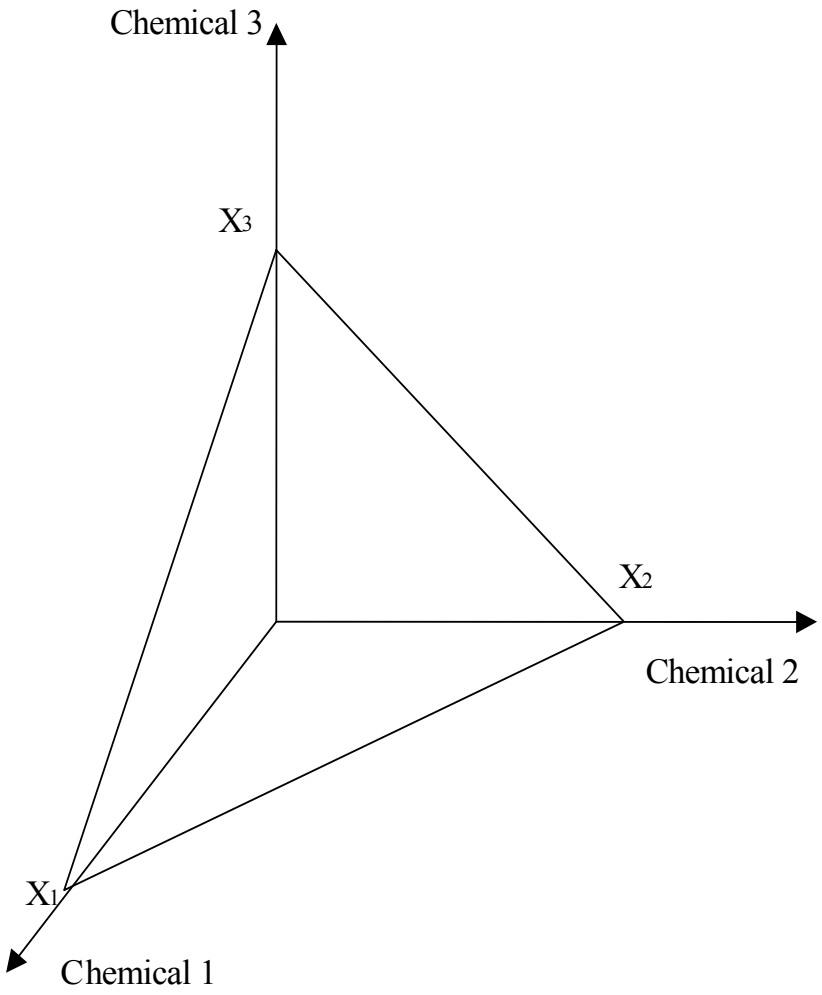
$$\sum_{i=1}^{c} \frac{x_i}{X_i} = 1.$$

A schematic of the resulting additivity surface for three chemicals is provided in Figure 1.

The additivity surface can be used to test the null hypothesis that the chemicals act in an additive fashion. The experimental data necessary and sufficient to support the estimation of the additivity surface are single chemical dose-response data (e.g., Gennings et al., 1997). To compare the response of an observed mixture of particular interest, such a mixture of the chemicals is also included in the experimental design. Therefore, the additivity surface (estimated using the single chemical data) is used to predict the response under additivity at the mixture point of interest. Comparison of the observed experimental response and the model-predicted response is made using a prediction interval. If the observed mixture response falls outside of the prediction interval, then the null hypothesis is rejected and we can say that either a greater than additive (synergistic) or less than additive (antagonistic) association exists. Conversely, if the observed experimental response falls within the prediction interval, the assumption of additivity is not rejected, providing evidence that dose additivity is a reasonable assumption for risk assessment. This general approach is described in detail by many authors, including Berenbaum (1985), Kelly and Rice (1990), and Gennings and Carter (1995).

Many researchers support the premise (Abelson, 1994) that human beings and experimental animals have repair mechanisms that respond to exposures to toxic insults in the environment.

Figure 1: Schematic of a three-dimensional additivity surface. The vertices are the points X_1, X_2, X_3 representing the concentration of each chemical alone which yields a fixed response. If the vertices of the surface correspond to the threshold parameters, then the surface is termed an 'additivity threshold surface'.



Such a premise suggests the existence of thresholds. A threshold is an exposure level below which any response is not distinguishable from background, while levels above the threshold result in a "dose-response" trend. The objective of this report is to demonstrate the use of a threshold additivity model in detecting and characterizing departure from additivity for mixtures of DBPs. The methodology has the advantage of being adaptable to large numbers of chemicals in combination with a reasonable sample size/experimental design. The required design includes single chemical concentration-effect data and only the combination point(s) (i.e., specific mixtures) of interest. This offers considerable advantages to examining the interactions of mixtures of chemicals formed during disinfection of drinking water as traditional full-factorial designs are likely to contain many mixture points that are of relatively little interest. Mixture points of most interest are those that mimic or are closely related to the relative proportions of the disinfection byproducts formed under given disinfection scenarios (a source water with certain humic acid and bromide characteristics, pH, temperature, and a particular disinfection scheme). Thus, limited and expensive laboratory resources can be focused on environmentally relevant mixtures.

The CD-1 mice data in the first example used to illustrate the methodology have been published previously (Gennings et al., 1997). Excerpts from the earlier report, including figures and tables, are included here. The mixture selected for study is that of four trihalomethanes formed when water is chlorinated. The four components are bromodichloromethane (BDCM), chlorodibromomethane (CDBM), chloroform ($CHCl_3$) and bromoform ($CHBr_3$). Table 1 shows an example of some individual mouse data that were generated and are illustrative of the effects at the dose-group level. Indicators of hepatic toxicity included increases in serum sorbitol dehydrogenase (SDH), alanine aminotransferase (ALT) and aspartate aminotransferase (AST), and increases in percent relative liver weight, supported by pathology results that indicate centrilobular necrosis. The endpoint used for the analysis was dose-related increases in SDH levels, representing the most sensitive endpoint in female CD-1 mice after 14 days of exposure by oral gavage in an aqueous vehicle.

The second example of this statistical methodology illustrates developmental toxicity for binary mixtures of $CHCl_3$ and $CHBr_3$ in the medaka fish. Concentration-effect data are available for $CHCl_3$ alone, $CHBr_3$ alone and mixtures of $CHCl_3$ and $CHBr_3$ where the endpoints are medaka embryo mortality at 24, 72 or 96 hours and the incidence of a heart or brain defect at these time points. The analysis in the report utilizes a composite binary response of death or defect at 96 hours.

Methods

The mice and experimental conditions for the CD-1 mice study were described by Gennings et al., (1997). These data, along with the medaka results, are also shown in a 1998 workshop report that provides details on this DBP mixtures research (U.S. EPA, 1998). The medaka data are also included in another paper in this report (Hartley et al., 1999, Section 4).

Table 1

Example Hepatotoxicity Data from THM Mice Experiments: Bromodichloromethane

Mouse ID#	Dose (mmol/kg/day)	ALT (IU/l)	AST (IU/l)	SDH (IU/l)	Body Weight Chg (g)	% Relative Liver Weight	Pathology Results
8	0	21	64	18	-1.8	5.12	Nothing detected
10	0.152	32	69	26	-1.1	4.01	Mild to moderate centrilobular vacuolation
66	0.305	32	47	28	6.6	3.97	Mild centrilobular vacuolation
52	0.76	35	66	33	2.1	5.27	Moderate centrilobular necrosis Moderate centrilobular vacuolation Moderate zone 2 vacuolation
83	1.52	476	249	403	-0.9	6.0	Moderate centrilobular necrosis Mild to moderate ballooned cell necrosis
81	3.05	277	242	249	-0.5	7.26	Severe centrilobular necrosis Moderate to severe ballooned cell necrosis

Briefly, female CD-1 mice (Charles River Breeding Laboratory, Raleigh, NC) were used. As stated in the Introduction, single-chemical dose response curves are needed to estimate the response surface under dose addition. To this end, each THM was assessed in a separate experiment. Groups of 8-20 mice were gavaged daily, for 14 days, with one of the four THMs, in 10% Alkamuls EL-620 at dosages of 0, 0.152, 0.305, 0.76 and 3.05 mmol/kg/day. Animals were gavaged in the morning with dosing solutions that were made fresh daily immediately prior to dosing. The gavage volume was 10 ml/kg. On the morning following the 14^{th} day of dosing, anoxia was induced by carbon dioxide. Serum was prepared and stored at -80^{o} C (nominal) until analyzed for the activity of sorbitol dehydrogenase (SDH) by automated procedures (COBAS Fara II) with appropriate reagents and standards. The mixture was tested as described for the single chemicals. The concentrations of the chemicals in the mixture were based on the average seasonal proportions at 35 water treatment facilities (Krasner et al., 1989). The tested mixture contained 0.208 mmol BDCM/kg, 0.084 mmol CDBM/kg, 0.569 mmol $CHCl_3$/kg and 0.012 mmol $CHBr_3$/kg for a total dosage of 0.872 mmol/kg/day. The experiment that tested the mixture contained a positive control for each of the four chemicals, for comparison with the single-chemical dose response curves used to generate the predicted response.

A culture of Japanese medaka (*Oryzias latipes*) was maintained under optimal conditions (Kirchen and West, 1976; Hartley et al., 1995) to obtain stage-specific embryos. Embryos were obtained in clusters, unfertilized embryos discarded, and chorionic filaments removed. Embryos were assigned randomly to individual glass vials with 5 mL of embryo-rearing solution. There were 10 embryos in the early high blastula stage per concentration group. DBP stock solutions were made up in DMSO and micropipetted into the appropriate exposure vial. There were daily volume exchanges and daily observation of lethality as well as neurological and circulatory system developmental abnormalities including heart development. As reported in U.S. EPA (1998) the binary mixtures experiment with $CHCl_3$ and $CHBr_3$ was designed based on the single-chemical dose response curves. A mixtures experiment with four (1:1 ratio) $CHCl_3$:$CHBr_3$ combination points was conducted to detect deviation from an additive toxicological response. The embryos were exposed to the following four binary mixtures of $CHCl_3$:$CHBr_3$: 10:10 mg/L, 15:15 mg/L, 20/20 mg/L and 25:25 mg/L. In this experimental design, two types of controls are required. They are the vehicle control (DMSO), and positive controls for $CHCl_3$ (75 mg/L) and $CHBr_3$ (25 mg/L).

Threshold additivity model for continuous responses

Let y_{ijk} be the response from the k^{th} experimental unit at the j^{th} dose group of the i^{th} chemical, $i=1,\ldots,c$; $j=1,\ldots,d_i$; $k=1,\ldots,n_{ij}$. The threshold additivity model using a power link function considered by Gennings et al. (1997) was

(1)

$$\mu^{\lambda} = \begin{Bmatrix} \beta_0, & if \sum_{i=1}^{c} \beta_i x_i < \delta \\ \beta_0 + \sum_{i=1}^{c} \beta_i x_i - \delta, & if \sum_{i=1}^{c} \beta_i x_i \geq \delta \end{Bmatrix}$$

where μ is the mean response, E(Y), when exposed to the chemical combination $\mathbf{X}=[x_1, x_2,\ldots,x_c]$. Using this model, the threshold for the i^{th} chemical alone is given by $\delta_i^* = \delta / \beta_i$, i=1,…,c. For the first example, c=4 as there were four chemicals under study. We assume that the variance of the response when exposed to X is var(Y)= τV(μ), where V(μ) is an assumed known function of the mean and τ is an unknown scale parameter. For the data described herein, we assume V(μ)=μ.

Threshold additivity model for binary responses

Concentration-effect data are available for chloroform (0, 25, 50, 75, 100, 125 ppm) and bromoform (0, 10, 25, 40, 50, 60 ppm) with the composite binary response of death or defect at 96 hours.

The probability of a response (defined as death or defect at 96 hours) was considered using a threshold model for binary responses with a logistic model above the threshold. That is, let Y_{ij} be the number of responses for the i^{th} chemical at the j^{th} dose group, i=1,2; j=1,…,6. The probability of a response under the assumption of additivity (i.e., no interaction) was modeled as:

(2)

$$\mu = \begin{Bmatrix} 1/(1+\exp(-\beta_0)), & \text{if} \sum_{i=1}^{2} \beta_i x_i < \delta_i \\ 1/(1+\exp(-(\beta_0 + \sum_{i=1}^{2} \beta_i x_i - \delta_i))), & \text{if} \sum_{i=1}^{2} \beta_i x_i \geq \delta_i \end{Bmatrix}$$

where, for this example,

x_1 is the concentration of $CHCl_3$ (ppm) in the total concentration given,
x_2 is the concentration of $CHBr_3$ (ppm) in the total concentration given,
β_0 is an unknown parameter associated with the background response rate,
β_1 is an unknown parameter associated with the effect of $CHCl_3$,
β_2 is an unknown parameter associated with the effect of $CHBr_3$,
δ_1 is an unknown parameter associated with the threshold for $CHCl_3$,
δ_2 is an unknown parameter associated with the threshold for $CHBr_3$,
Var(Y) is assumed to be τμ(1-μ).

For both examples, model parameter estimates were found using the method of maximum quasi-likelihood (e.g., McCullagh and Nelder, 1989). As the models given in (1) and (2) are not smooth at the join points given by δ_i, the usual linearization methods may have difficulty converging to global parameter estimates. Therefore, we employ the Nelder Mead direct search algorithm (Nelder and Mead, 1965). For the model given in (1), the algorithm conditions on the value of the power parameter, l. Reasonable values of l are found by considering a range of values and selecting the one with the most reasonable interpretation.

The initial scheme in a threshold analysis is to investigate the significance of the dose-response relationship. A quasi-likelihood ratio test statistic is used to compare the full model given in (1) to a reduced model given by $\mu^{\lambda}=\beta_0$ under the restriction that $\beta_1=\ldots=\beta_c=\delta=0$. The statistic is compared to $\chi^2_{(1-\alpha,c+1)}$, or more conservatively, $F_{(1-\alpha;c+1,\ N-c-2)}$ (see Gennings et al., 1997). If the hypothesis of no dose-response trend is rejected, a confidence interval on each of the single chemical threshold parameters, δ_i^*, can be constructed by

$$\hat{\delta}_i^* \pm t_{1-\alpha/2;N-c-2}\sqrt{Var(\hat{\delta}_i^*)} \tag{3}$$

where the form of $Var(\hat{\delta}_i^*)$ is described in Gennings et al. (1997).

To compare the observed response at **X** to that predicted under the hypothesis of additivity, a prediction interval is used. Let $\bar{y}_{\mathbf{X}}$ be the observed sample mean response at the combination **X** with sample variance var($\bar{y}_{\mathbf{X}}$). Denote the estimated predicted response using the model given in (1) by $\hat{y}_{\mathbf{X}}$ with estimated variance var($\hat{y}_{\mathbf{X}}$) with details provided in Gennings et al. (1997). Then a 100(1-α)% prediction interval for the sample mean response at **X** under the assumption of additivity is given by

$$\hat{y}_{\mathbf{X}} \pm t_{(1-\alpha/2;N-c-2)}\sqrt{Var(\hat{y}_{\mathbf{X}})+Var(\bar{y}_{\mathbf{X}})} \tag{4}$$

If the observed sample mean, $\bar{y}_{\mathbf{X}}$, is not included in the prediction interval, then it is reasonable to conclude that the data do not support the assumption of additivity. If the observed sample mean is less extreme than the prediction interval, then antagonism can be claimed at **X**; if the observed sample mean is more extreme than the prediction interval, then synergism can be claimed at **X**.

Results

CD-1 Mice Example

The purpose of the CD-1 mouse illustration is to determine if a combination of the four chemicals (BDCM, CDBM, $CHCl_3$ and $CHBr_3$) results in a response different from an additive one as predicted by the additivity model given in (1). SDH was selected as the response endpoint. The doses of each chemical in the single chemical experiments were 0, 0.152, 0.305, 0.76, 1.52 and 3.05 mmoles/kg/day. The resulting sample means, standard deviations, and sample sizes are provided in Gennings et al. (1997).

The threshold additivity model given in (1) was fitted to the data using the method of maximum quasi-likelihood. The power function on the mean was included to facilitate a curvilinear dose-response relationship above the threshold surface. A value of the power transformation of 0.5 was used in the analysis because it was roughly associated with the highest quasi-likelihood value over a grid of power parameters. In addition, the variance was assumed to be of the form $Var(Y)=\tau\mu$. A Nelder-Mead algorithm was used to estimate the unknown parameters given in (1). Graphs of the fitted dose-response curves are given in Figure 2.

The quasi-likelihood ratio test of no dose-response for any of the chemicals was rejected ($p<0.001$). The 95% confidence intervals on the threshold for each chemical (equation 3) are provided in Table 2. All four intervals include zero, thereby indicating a significant threshold was not found for any of the chemicals.

The mixture point of interest for the four chemicals was based on their average seasonal proportions of 35 water treatment facilities as reported by Krasner et al. (1989). The proportion (in mmoles) is (0.24: 0.10: 0.65: 0.01) for (BDCM: CDBM: $CHCl_3$: $CHBr_3$). The actual dose combination considered was (0.208, 0.084, 0.568, 0.012) mmoles/kg/day. This point is referred to as **X**.

Using the model given in (1) and the estimated model parameters, the predicted SDH response under additivity at **X** is $\hat{y}_{\mathbf{X}} = 42.9$ with standard error of 2.79. The observed mean SDH response was $\bar{y}_{\mathbf{X}} = 43.9$ with standard deviation 15.8. Thus $Var(\bar{y}_{\mathbf{X}})$ is 13.1 with n=19. Using equation (4), a 95% prediction interval on the response at **X** under the assumption of additivity is [33.9, 51.9]. This prediction interval is plotted (denoted by '+') with the observed mean response (denoted by the square) along the predicted response under additivity of the mixture ratio ray of (0.24: 0.10: 0.65: 0.01) in Figure 3. Since the observed mean response at **X** is included within the prediction interval, these data provide no evidence of departure from additivity at the specified combination of the four chemicals.

Figure 2a: Observed SDH responses (*) and model-predicted mean responses based on the threshold model given in (1) for doses of BDCM alone.

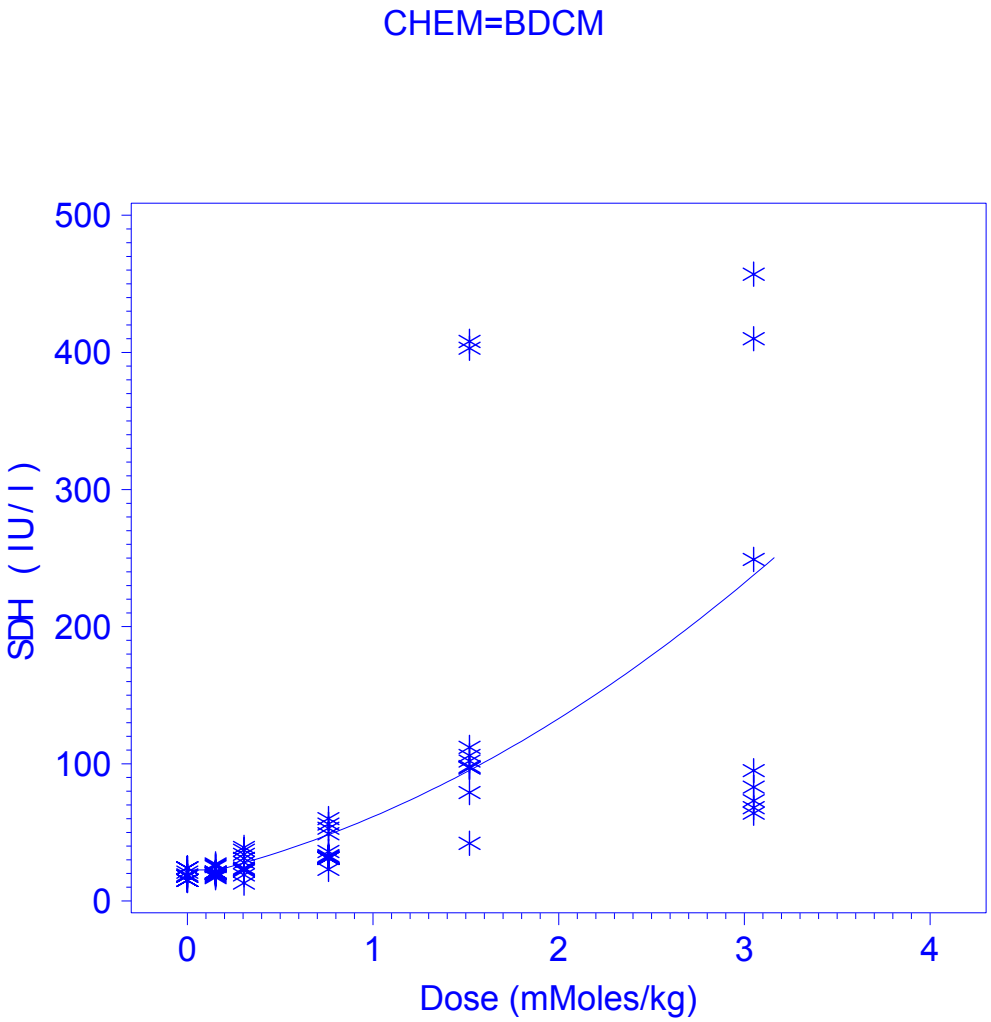


Figure 2b: Observed SDH responses (*) and model-predicted mean responses based on the threshold model given in (1) for doses of CDBM alone.

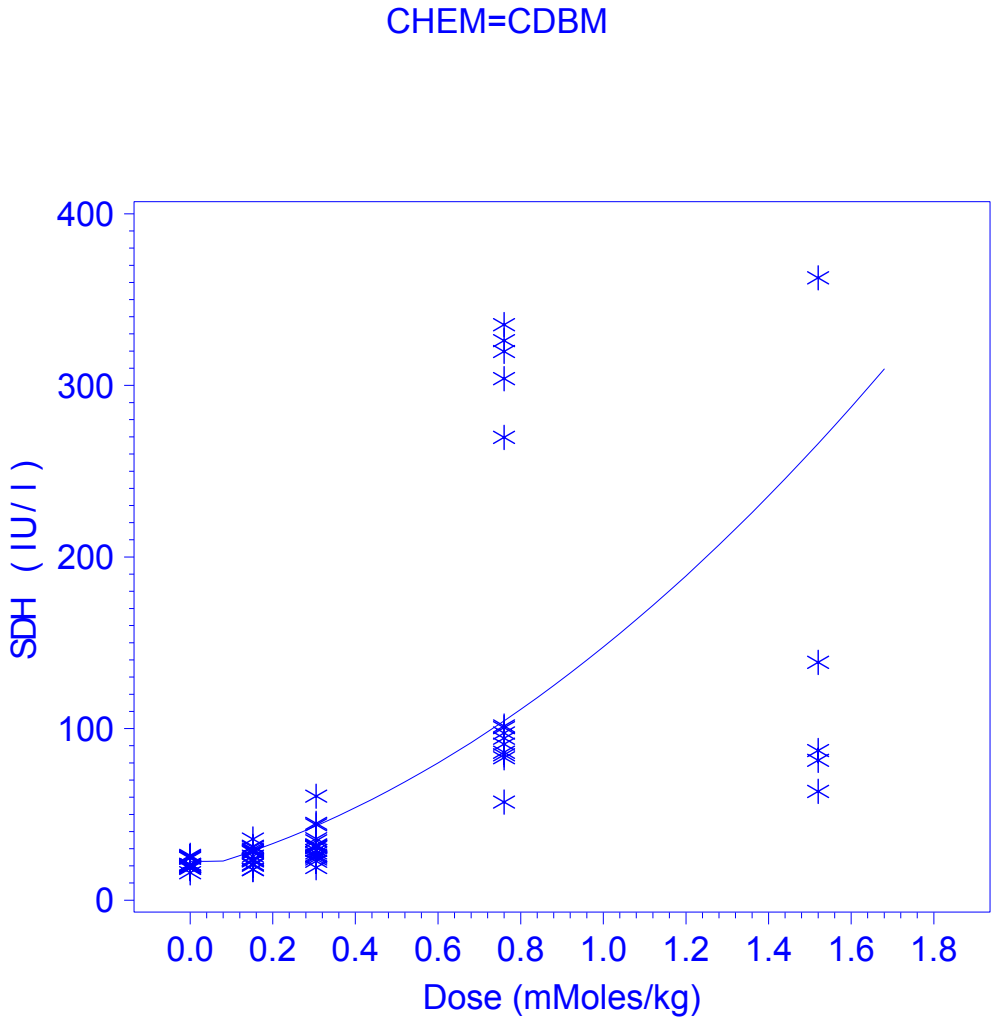


Figure 2c: Observed SDH responses (*) and model-predicted mean responses based on the threshold model given in (1) for doses of CHCl3 alone.

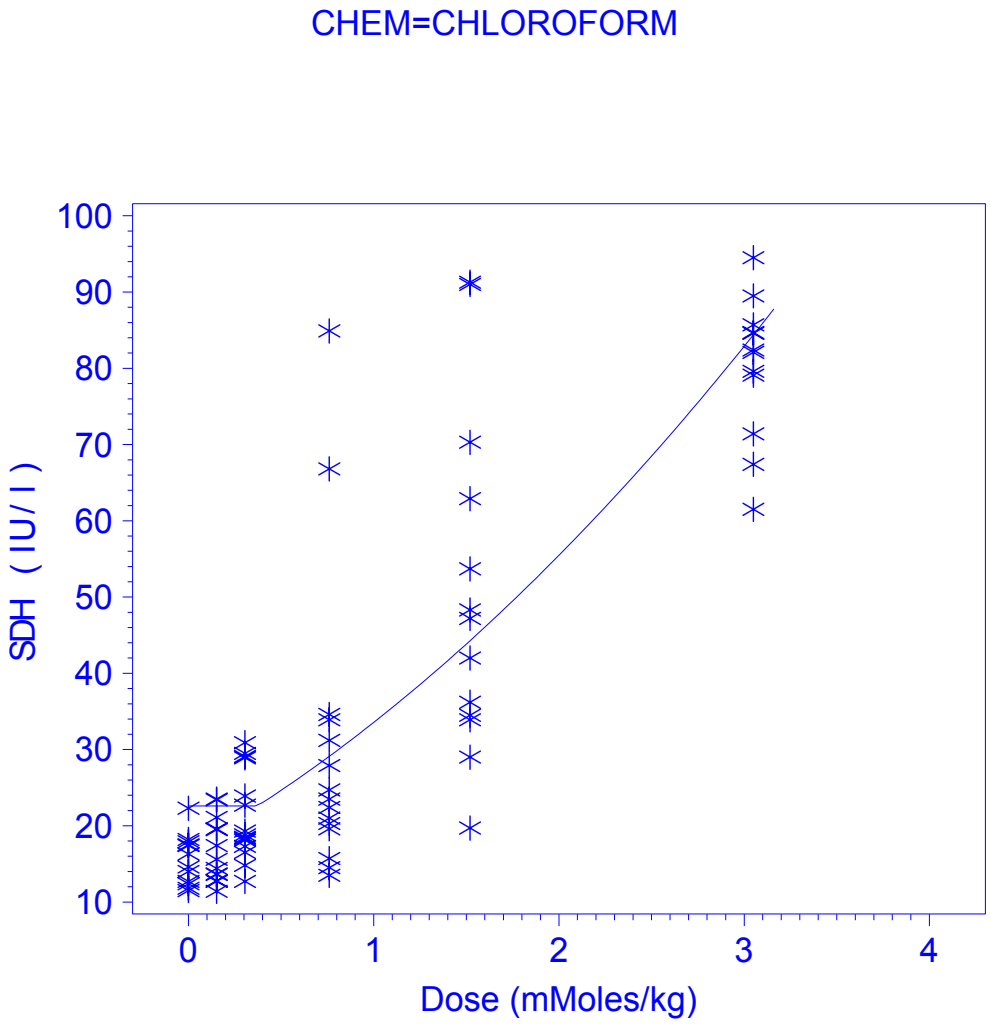


Figure 2d: Observed SDH responses (*) and model-predicted mean responses based on the threshold model given in (1) for doses of CHBr3 alone.

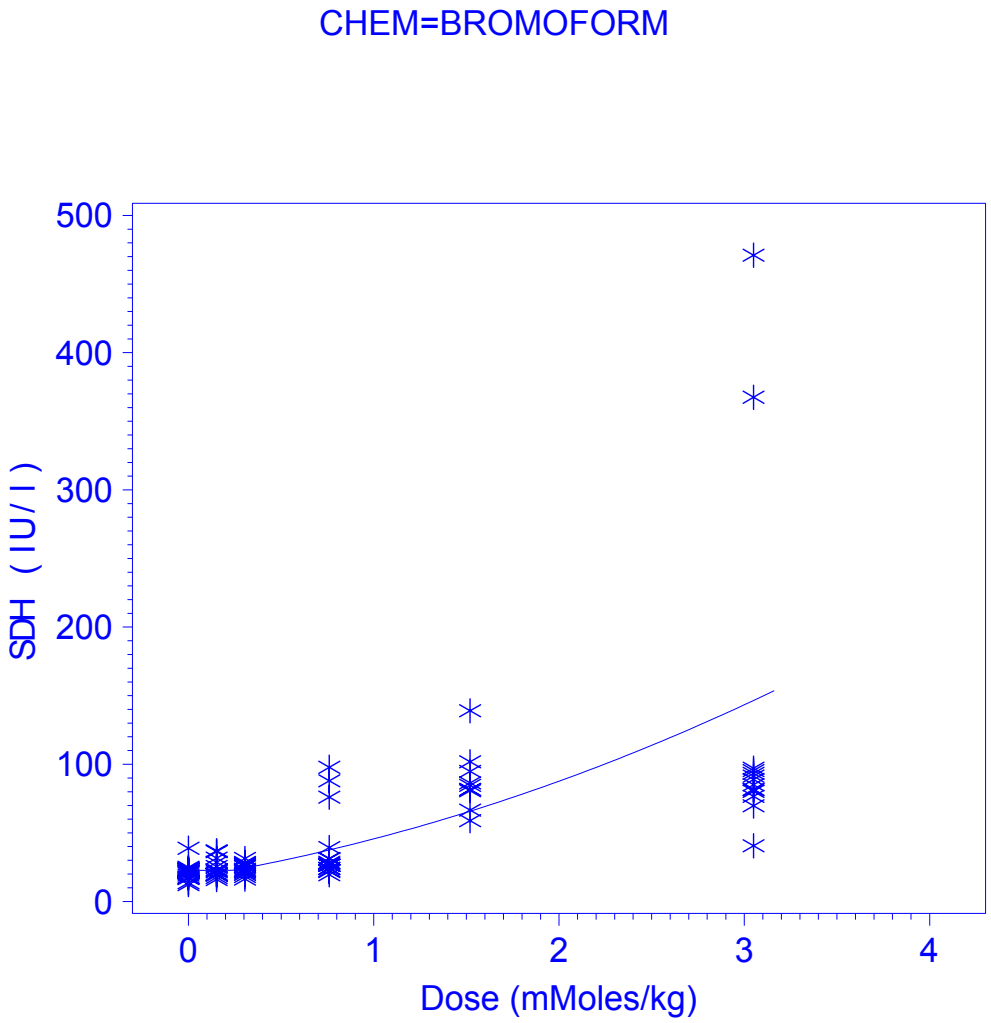
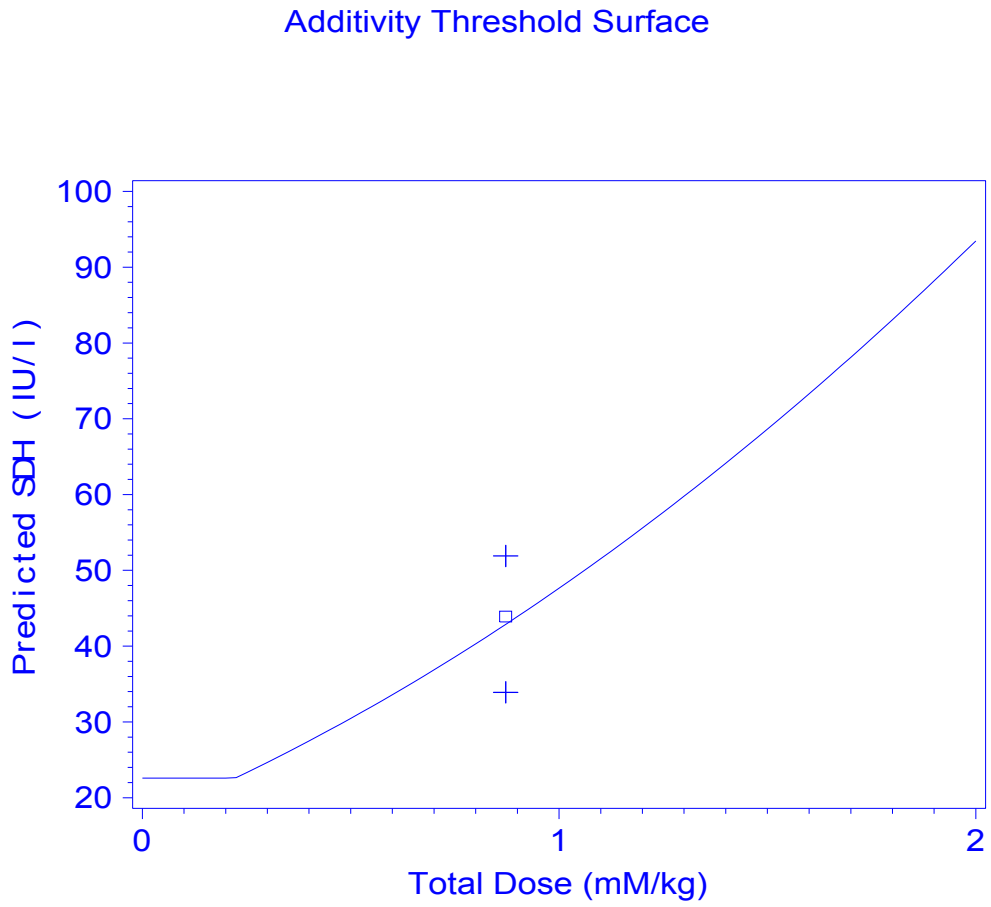


Table 2

Using the parameter estimates provided in Table 2 of Gennings et al. (1997), the threshold estimates and resulting 95% confidence intervals were obtained as in (2). The table is reproduced from Table 3 in Gennings et al. (1997).

Chemical	Threshold Estimate (SDH, IU/l)	95% Confidence Interval
BDCM	0.167	[0, 0.364]
CDBM	0.077	[0, 0.168]
$CHCl_3$	0.372	[0, 0.799]
$CHBr_2$	0.236	[0, 0.512]

Figure 3: Using the threshold additivity model in (1) and the mixing ratio (in mmoles) of (0.24: 0.10: 0.65: 0.01) for (BDCM: CDBM: CHCl3: CHBr3), the predicted response under additivity is plotted as a line. The observed sample mean at the total dose of 0.872 mmoles is plotted as a square. The 95% prediction limits under additivity are denoted as '+'.



Medaka Fish Example

The LC50 for $CHCl_3$ was determined in 10 embryos/concentration exposed to 0, 25, 50, 75, 100 or 125 ppm $CHCl_3$ for 96 hours in aquatic medium. The LC50 was determined to be 59 ppm (95% CL= 50-72 ppm). There was a dose-related increase in combined mortality and severe neurological and circulatory defects with incidence reaching 100% at >100 ppm. Heartbeat rate was measured in surviving embryos at 72 and 96 hours. At 72 hours, mean heartbeat rate decreased with increasing concentration in all dose groups relative to controls. The heartbeat rate was statistically significantly reduced in the two highest dose groups (100 and 125 ppm) relative to controls. At 96 hours, heartbeat rate was significantly decreased relative to controls in all groups with the exception of the low (25 ppm) and high concentrations (125 ppm). Mean heartbeat rate was statistically significantly reduced in the 50, 75, and 100 ppm groups, but not in the high concentration group. (A possible explanation for this result is hardy survivors and the fact that there were only three embryos left alive in the 125 ppm concentration group to contribute data to the analysis.)

The LC50 for $CHBr_3$ was determined in 10 embryos/concentration exposed to 0, 10, 25, 40, 50 or 60 ppm $CHBr_3$ for 96 hours in aquatic medium. The LC50 was determined to be 15.7 ppm (95% CL= 6-21 ppm). There was an increase in combined mortality and severe neurological and circulatory defects in all exposed groups. At 10 to 25 ppm, the incidence for these effects was 50-60%, whereas the incidence reached 100% at $\geq$40 ppm. Heartbeat rate was measured in surviving embryos at 72 and 96 hours yielding inconsistent results. At 72 hours, mean heartbeat rate was statistically significantly increased at 10 ppm and statistically significantly decreased at 50 ppm relative to controls; other concentration groups had heartbeat rates comparable to controls. At 96 hours, heartbeat rate was statistically significantly increased at 10 and 25 ppm, statistically significantly decreased at 60 ppm, and comparable to controls at other concentrations.

In the experiment studying the mixture containing $CHCl_3$ and $CHBr_3$, mortality in the vehicle control, and $CHCl_3$ positive control and $CHBr_3$ positive control was 10, 30, and 20%, respectively. There was no mortality in any of the four binary mixture groups. When the incidence of mortality was combined with the incidence for severe neurological or circulatory defects, the combined incidence was 20, 40, 60, 40, 30, 40 and 20% for the control, $CHCl_3$ alone, $CHBr_3$ alone, 10:10, 15:15, 20:20 and 25:25 ppm groups, respectively (Table 3). The type of neurological and developmental defects observed in the mixture groups were similar to those observed in the single chemical groups.

Response incidences observed for the positive controls in the mixture experiment were not different from those observed in the single-chemical experiment. While the single chemical experiments showed statistically significant dose-response trends, the estimates of the thresholds were extrapolations. A smooth model (i.e., a non-threshold model, here a logistic model) was

Table 3

Observed and model predicted responses under additivity using a logistic model for chloroform: bromoform mixtures in medaka fish.

	Observed Combined Incidence of Death/Defect	Predicted Death/Defect Under Additivity	95% Prediction Interval
Controls			
Control	0.2	0.024	[0.001, 0.370]
$CHCl_3$: 75 ppm	0.4	0.607	[0.248, 0.878]
$CHBr_3$: 25 ppm	0.6	0.669	[0.350, 0.884]
$CHCl_3$: $CHBr_3$			
10:10 ppm	0.4	0.198	[0.042, 0.583]
15:15 ppm	0.3	0.442	[0.161, 0.764]
20:20 ppm	0.4	0.716	[0.341, 0.925]
25:25 ppm	0.2	0.890	[0.461, 0.987]

suggested for modeling the probability of death or severe defect to avoid over-parameterization of the threshold additivity model. Departure from additivity (antagonism) was detected in the highest total dose group (25:25 ppm) because the observed response (0.2) fell below the 95% prediction interval (0.461, 0.987). There was no statistically significant departure from additivity at the three lower mixture groups (Table 3).

Discussion

Important features of the methodology illustrated in this report include its flexibility in detecting regions of additivity and/or departure from additivity while maintaining an experimental design that is practical in size. Based on further work from our laboratory, we recommend that the experimental design include dose-response data for each component in the mixture. Combination points of interest can be selected based on relevant mixtures (e.g., the Krasner mixture is similar to the median mixture of the 35 water treatment facilities). The logic of the procedure is to predict under the hypothesis of additivity at the particular mixture(s) of interest. If the observed response exceeds that predicted under additivity, a synergism is claimed at the particular mixture; if the observed response is less than that predicted under additivity, an antagonism is claimed; otherwise, departure from additivity cannot be claimed. This comparison is based on a prediction interval as it takes into account the variability in the predicted response as well as the observed response. Additionally, results of the threshold additivity model may be considered in a biological and mechanistic context. The results may be consistent with our

present understanding of the mechanism(s) underlying the toxicity of the single chemical or, alternatively, might suggest that other mechanistic hypotheses may need to be explored.

To establish or refute dose additivity as the mode of joint action for DBPs is an important concept for the development of risk assessment methods applicable to these chemicals. A key question for those interested in drinking water treatment technologies is whether differences exist in human health consequences among the various options available to disinfect water. Dose additivity is the basic assumption for many of the current multiple chemical risk assessment methods that can be used for the estimation of health effects, given exposure to a particular treatment train and source water characteristics. Thus, further development of mixture toxicity data for DBPs and statistical methods that can provide information on this joint action are needed for the evaluation of treatment options.

References

Abelson, P. (1994) Editorial in *Science*, September, 265: 1507.

Berenbaum, M.C. (1985) The expected effect of a combination of agents: The general solution, *Journal Theoretical Biology*, 114: 413-431.

Gennings, C., and Carter, W.H., Jr. (1995) Utilizing concentration-response data from individual components to detect statistically significant departure from additivity in chemical mixtures, *Biometrics,* 51: 1264-1277.

Gennings, C., Schwartz, P., Carter, W.H., Jr., Simmons, J.E. (1997) Detection of departures from additivity in mixtures of many chemicals with a threshold model, *Journal of Agricultural, Biological, and Environmental Statistics,* 2: 198-211.

Hartley, W.R., Thiyagarajah, A., and Mizell, M. (1995) Ectopic hermaphroditic gonad in the cranial cavity of Japanese medaka, *Journal of Aquatic Animal Health*, 7: 172-177.

Kelly, C., and Rice, J. (1990) Monotone smoothing with application to dose-response curves and the assessment of synergism, *Biometrics*, 46: 1071-1085.

Kirchen, R., and West, W. (1976) The Japanese medaka its care and development, Carolina Biological Supply Co, Burlington, North Carolina.

Krasner, S.W., McGuire, M.J., Jacangelo, J.G., Patania, N.L., Reagan, K.M., Aieta, E.M. (1989) The occurrence of disinfection by-products in U.S. Drinking Water, *Journal American Water Works Association*, 81: 41-53.

Loewe, S. (1953) The problem of synergism and antagonism of combined drugs, *Arzneimittle forshung*, 3: 285-290.

Loewe, S. and Muischnek, H. (1926) Uber kmobinationswirkunger. I. Mitteilung: Hiltsmittel der fragstellung. Naunyn-Schmiedebergs, *Arch. Pharmacol*, 114: 313-326.

McCullagh, P., and Nelder, J.A. (1989) *Generalized Linear Models (2nd ed)*, New York: Chapman and Hall.

Nelder, J.A., and Mead, R. (1965) A simplex method for function minimization, *The Computer Journal,* 7: 308-313.

U.S. EPA (1998). Workshop Report: The risk assessment of mixtures of disinfection by-products (DBPs) in drinking water. National Center for Environmental Assessment, Cincinnati, OH. NCEA-C-0459.

PROGRESS REPORT: CARCINOGENICITY OF BENZENE IN THE JAPANESE MEDAKA (*ORYZIAS LATIPES*) EXPOSED DURING EMBRYOLOGICAL DEVELOPMENT

William R. Hartley, Sc.D. and A. Thiyagarajah, Ph.D.

Tulane University
School of Public Health and Tropical Medicine
Environmental Diseases Prevention Research Center
New Orleans, LA

September 1999

Abstract

Benzene is a volatile organic compound, found in contaminated air, water, sediment, soils, food, and tobacco smoke. Benzene is classified as a known human carcinogen by U.S. EPA and has produced an array of tumor types in rodent studies. The goals of this study were to determine the carcinogenic effects of benzene in Japanese medaka using a 10-day embryonic exposure methodology, to contrast the results from the standpoint of cross-species comparisons, and to discuss potential use of these methods for risk assessment. In this exploratory research, the basis for limiting exposure periods to the embryo were the known sensitivity of the medaka to the carcinogens in long-term exposure; transparency of the chorionic membrane (membrane surrounding the developing embryos) such that developmental effects could be continuously or periodically observed; simplicity of exposure regimens; and the ability to focus on critical periods of development. Medaka embryos were exposed to 0, 43.5, 100, and 150 mg/L benzene beginning at early high blastula for 10 continuous days.

The only liver tumors noted in the one-year-old medaka exposed as embryos to benzene were hepatic adenoma and cholangioma. There was an increasing trend ($p \leq 0.1$) in the total tumor incidence in fish surviving after one year in the two high concentration (100 and 150 mg/L benzene) groups, with one case of hepatic cholangioma in the control group. Based on the sensitivity of the medaka to known carcinogens in lifetime (chronic) exposure studies and frequent use in cancer research (Hawkins et al., 1985; Chen et al., 1996), this study used an embryo exposure route to determine the sensitivity of the medaka to the carcinogenicity of benzene. Because the exposure period was limited to 10 days of embryogenesis and new methods were being used, high (shock) concentrations of benzene of known embryo lethality were administered to optimize penetration of benzene through the chorion. The hepatocellular adenomas and cholangiomas observed in this study were benign neoplasms. The malignant neoplasms associated with benzene exposure in the previously summarized rodent literature were not observed in this study. Considering the marginal statistical significance of the increased liver tumors and the absence of malignant liver tumors, this medaka embryo exposure methodology does not indicate the known carcinogenic potential of benzene and does not reflect the highly variable carcinogenic response reported in numerous rodent bioassays. The low incidence of non-neoplastic proliferative and degenerative lesions observed in this study are probably a consequence of normal aging. To further evaluate the utility of the embryo exposure methodology for risk assessment, a similar protocol will be tested for the potent animal carcinogen diethylnitrosamine (DEN). The exposure period may also be extended to include young fish in the range of several weeks post hatch.

Introduction

The purpose of this study was to determine the carcinogenic effects of benzene in Japanese medaka using a 10-day embryonic exposure methodology. Another goal was to contrast our results from the standpoint of cross-species comparisons and discuss potential use of our methods for risk assessment. This is the first of a series of ongoing experiments with other chemicals including drinking water disinfection by-products and the animal carcinogen, diethylnitrosoamine (DEN).

In this exploratory research, the basis for considering limiting exposure periods to the embryo were: known sensitivity of the medaka to the carcinogens in long-term exposure; transparency of the chorionic membrane (membrane surrounding the developing embryos) such that developmental effects could be continuously or periodically observed; simplicity of exposure regimens; and the ability to focus on critical periods of development. We hypothesize that studies in which the medaka embryo is used for cancer assessment and other chronic effects should include a continuous exposure period from fertilization to well after development of the liver rudiment or at least greater than 128 hours (>5 days). We exposed the embryos to benzene beginning at early high blastula for 10 continuous days. At day 10 the liver is well developed and conspicuous. It consists of globules and nearly covers the greenish urinary bladder on the left side of the embryo. The spleen is a crimson colored rudiment between the urinary and swim bladders. Most embryos hatch between the 11^{th} and 14^{th} day (at 25°C) of development and begin feeding within one day.

Benzene

Benzene is a volatile organic compound released to the environment by natural or human activities. Over one billion gallons of benzene is produced in the United States annually, often used as raw material for synthesis of a variety of chemicals (Purcell, 1978) and also as solvents. See ques in abstract. It is found in contaminated air, water, sediment, soils, food, and tobacco smoke. (Huff et al., 1989). Based on carcinogenicity studies, benzene is classified as a group A (known human carcinogen) by U.S. EPA (1999).

In evaluating the utility of limiting exposure to the critical periods of embryological development for risk assessment, our results must be contrasted with the extensive human and rodent carcinogenicity database on benzene. The unequivocal epidemiological data, particularly on carcinogenicity, make cross-species comparisons of the effects of benzene possible. Long-term exposure to benzene results in decreased red blood cells and anemia. Other primary non-carcinogenic effects include immune system suppression and increased risk of infectious disease. The primary chronic effect, unequivocally based on epidemiological studies (inhalation exposure), is increased risk of acute myeloid leukemia (AML) (Hurley et al., 1991; Hunting et al., 1995). In AML, there is a diminished production of normal erythrocytes, granulocytes, and platelets, which leads to death by anemia, infection and hemorrhage. It is assumed there is a similar risk in humans from oral exposure but this has not been proven by adequate studies.

Studies on the carcinogenicity of benzene in rodents, unlike epidemiological studies, report an array of increased exposure-related tumor types. In 15 to 16 week, 5 days/week, 4 to 7 hours/day inhalation exposure studies in the 100 to 300 ppm benzene range, significant increases in neoplasms were reported: hepatomas and Zymbal gland (auditory sebaceous glands) carcinomas in Sprague-Dawley rats (Maltoni et al., 1982a, 1983, 1985); leukemia in the CBA/Ca mouse (Cronkite, 1986); thymic and non-thymic lymphoma in the C57BL/6BNL mouse (Cronkite et al., 1984, 1985); lymphomas and myelogenous neoplasms, mylelogenous neoplasms, Zymbal gland carcinoma, squamous cell carcinoma and mammary carcinoma, lung papillary adenocarcinoma in the CBA/CaBNL mouse (Cronkite et al., 1989), and lymphoma in CBA/Ca mouse (Farris et al., 1993). In chronic inhalation studies (86 weeks to lifetime), 5 days/week, 4 to 7 hours/day inhalation exposure studies in the 100 to 300 ppm benzene range, significant increases in neoplasma were reported: hepatomas, Zymbal gland carcinoma, and myelogeneous leukemia in Sprague-Dawley rats (Maltoni et al., 1982a,b, 1983, 1985; Snyder et al., 1984), and hematopoetic neoplasms including thymic lymphomas in AKR/J,C57BL6J mouse (Snyder et al., 1978, 1980). In 52-week, 4-5 days/week oral exposures in the Sprague-Dawley rat, increased incidences of Zymbal gland carcinoma, oral carcinoma, fore-stomach tumors, and liver angiosarcoma in the 50-250 mg/kg/day benzene dose range were reported (Maltoni et al., 1983, 1985, 1989). In chronic (52 weeks to 2 years) oral studies in the 25 to 500 mg/kg/day benzene dose, increased tumor incidences reported include: oral cavity squamous cell papilloma and carcinoma, and Zymbal gland tumors in the F-344N rat (Huff et al., 1989) and Zymbal gland carcinoma, oral and nasal cavity carcinoma, and angiosarcoma of the liver in Sprague-Dawley and Wistar rat (Maltoni et al., 1983, 1989).

Chronic toxicity of benzene or benzene derivatives, but not carcinogenicity, have been studied in several fish species but not in medaka. Benzene has been shown to influence the rate of respiration in striped bass (Meyerhoff, 1975), chinook salmon (Broncksen and Biley, 1973), and in mullet (Correa and Garcia, 1990). Rainbow trout exposed to benzene exhibited dose-response related chromosomal aberrations (Al-Sabti, 1985).

Materials and Methods

Embryo Source

The culture of medaka was maintained under optimal conditions (Kirchen and West, 1979; Hartley et al., 1995) to obtain stage specific embryos. The broodstock of Japanese medaka fish was purchased from Carolina Biological Supply Company (Burlington, NC) and reared in a recirculating water system in our laboratory. The brood stock and experimental grow-out fish were maintained at 21 to 26°C and a photoperiod of 14 hours of artificial light and 10 hours of dark. Dechlorinated municipal (New Orleans) water was used in the recirculating fish tanks. Approximately 50% of the municipal water was further treated to remove potential trace organics, metals, and solids. The treatment system consisted of five units with activated carbon filtration, cation exchange, anion exchange, and two mixed media columns (Culligan, New Orleans, LA). The quality of treated water met the requirements for culture and toxicity testing of aquatic organisms (APHA et al., 1992). The tank maintenance included volume changes and

cleaning gravel beds (twice per month) and weekly or more frequent chemical analysis for water quality parameters including pH, ammonia, nitrate, and nitrite. Eggs were collected and reared according to methods described by Kirchen and West (1979). Fish were fed three times daily with Tetra Min Flake food or fry food (TetraWerke, Germany), and supplemented with live or frozen *Artemia* (brine shrimp) three times per week.

Test Chemical

High pressure liquid chromatography (HPLC) grade benzene (99.9%) was purchased from Sigma-Aldrich Company, St. Louis, MO. Stock solutions and exposure solutions of benzene were made up without carrier, because of its adequate solubility in embryo-rearing media (Carolina Biological Company, Burlington, NC). All solutions were kept tightly sealed to reduce volatilization. They were analyzed by gas chromatography (purge and trap), U.S. EPA Solid Waste 846 Method 9021. Solutions not within 10% of target concentrations were rejected.

Exposure

Eggs were collected from the female medaka, checked for viability, and then randomly assigned to each exposure chamber. Upon removal, the embryos were typically in the 16-cell stage. The embryos were exposed to benzene dissolved in embryo-rearing solution (ERS), raised to hatch in ERS (Rugh, 1962), and maintained at 25°C. Exposure typically began in the 64-cell stage to early high blastula. A total of 300 embryos were exposed in groups of 50 per chamber for 10 days to nominal concentrations of 0, 43.5, 100, and 150 mg/L of benzene without removal of chorionic filaments. Based on the Lethal Concentration (LC50) for benzene of 177 mg/L (95%CL 151-202), the exposure concentrations represented approximately 25%, 56%, and 85% of the LC50 value. High concentrations were used to optimize transport of the benzene across the chorion and ensure a high dose exposure. At the end of a 10-day exposure, surviving embryos were rinsed several times in ERS, transferred to hatching chambers, and monitored daily for hatching. All newly hatched fish were transferred to grow-out chambers. At one year, the surviving fish were then killed by an overdose of MS222 (Crescent Chemicals, AZ) and all organs examined for tumors. Fish were fixed whole in 10% neutral buffered formalin (NBF) after a mid-ventral incision. Fish were cut in half saggittally and then decalcified in Surgipath decalcifier I (Surgipath Medical Industries, Richmond, IL). Fish were processed routinely for histological examination. All major tissue types were examined for neoplasms.

Results and Discussion

The incidence of hepatic neoplastic, non-neoplastic proliferative, and degenerative lesions are summarized in Table 1. The were no other neoplasms in other tissues. Because the neoplasms were observed only in the liver, we have also included observations of non-neoplastic proliferative and degenerative lesions. The diagnostic criteria used to classify these lesions in

Table 1

Incidence of Hepatic Lesions in One-Year-Old Medaka Exposed to Benzene During Embryological Development

Liver Lesions	Benzene Concentrations			
	0 mg/L (n=59)	43.5 mg/L (n=31)	100 mg/L (n=44)	150 mg/L (n=39)
No lesions	42	20	27	16
Neoplasms				
Adenoma	0	1	1	1
Cholangioma	1	1	3	3
Total Tumors	1	2	4	4
Percent Total Tumors	1.7	6.5	9.1	10.3
Fisher Exact Test, p value		0.27	0.10	0.08
Non-Neoplastic Proliferative				
Basophilic foci	1	0	0	1
Eosinophilic foci	1	0	3	0
Clear-cell foci	5	4	7	7
Vacuolated cell foci	6	8	6	11
Degenerative				
Spongiosis hepatis	3	3	5	2
Necrosis	2	2	4	3

this study are published by the National Toxicology Program Pathology Working Group (Boorman et al., 1997). Figure 1 illustrates the histological appearance of a control medaka liver.

Neoplasms

Hepatic adenoma and cholangioma were the two types of neoplasms observed in this study. Hepatic adenomas (Figure 2) were distinct, small to moderately sized lesions clearly demarcated from the neighboring tissue. The cell populations were basophilic, polygonal cells, fairly uniform in size and shape. No mitotic figures were seen. Cholangiomas were well differentiated proliferating bile ducts with dilated lumen clustered in a focal area and expanding into the neighboring tissue. Proliferating epithelium was often surrounded by fibrous stroma. In two cases, proliferating bile duct epithelium was thrown into papillary projections and the duct lumen was abnormally dilated and cystic. Although histologically these two cases were diagnosed as papillary cystic cholangioma (Figure 3), in the summary table these were grouped under cholangioma because of the biliary cellular origin of this benign neoplasm.

Non-neoplastic Proliferative Lesions

Non-neoplastic proliferative lesions observed were basophilic, eosinophilic, and clear-cell foci. Incidence of these lesions is summarized in Table 1. Basophilic foci (Figure 4) consisted of hepatocytes smaller than the neighboring hepatocytes and containing basophilic cytoplasm. Cells were polygonal in shape with no nuclear irregularity. Mitotic figures were not common. The margin of the lesion was often irregular. Eosinophilic foci (Figure 5) consisted of polygonal hepatocytes were larger than the neighboring hepatocytes, and contained eosinophilc cytoplasm. The margin of the lesions was irregular. No mitotic figures were seen. Clear-cell foci (Figure 6) consisted of large polygonal hepatocytes containing sparce eosinophilc granular cytoplasm that often gave the appearance of ground glass. The nuclei were often centrally located. Occasionally, eccentric nuclei were also seen in these lesions, which often had irregular margins.

Degenerative Lesions

Spongiosis hepatis

Spongiosis hepatis (Figure 7) occurred in both control and benzene-exposed medaka . This lesion was characterized by multilocular spaces filled with pale eosinophilic material. These cystic spaces were lined by perisinusoidal cells. This lesion was multifocal in all medaka examined in this study. The size varied from small to extensive.

Vacuolation of hepatocytes occurred either diffusely or in foci . These vacuoles were often single, large, and spherical. Focal necrosis (sometimes multifocal) occurred in both control and exposed fish. Focal areas of hyalinization of hepatocytes occurred in two fish exposed to 150 mg/L benzene, but not in any other group of fish.

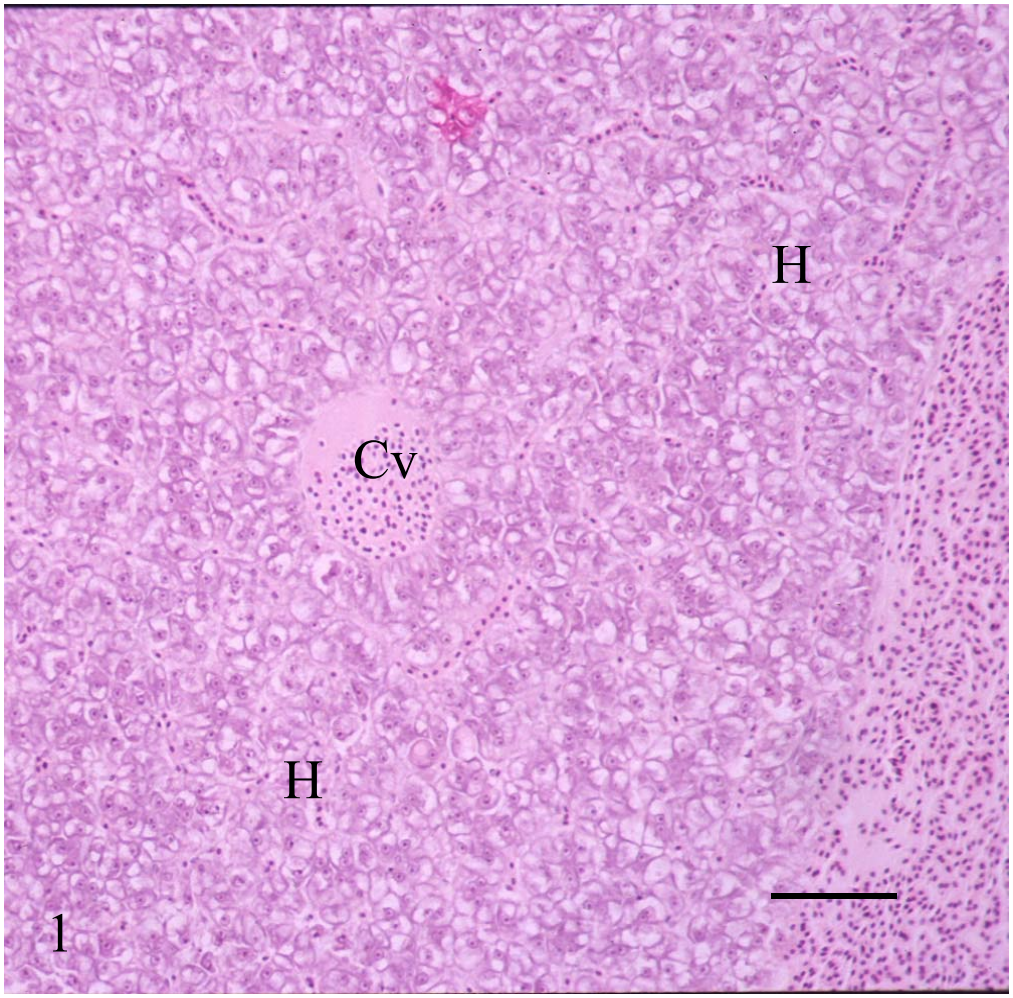


Figure 1. Histomorphology of a control medaka liver showing hepatocytes (H) and central vein (Cv). H&E, bar=70 μm.

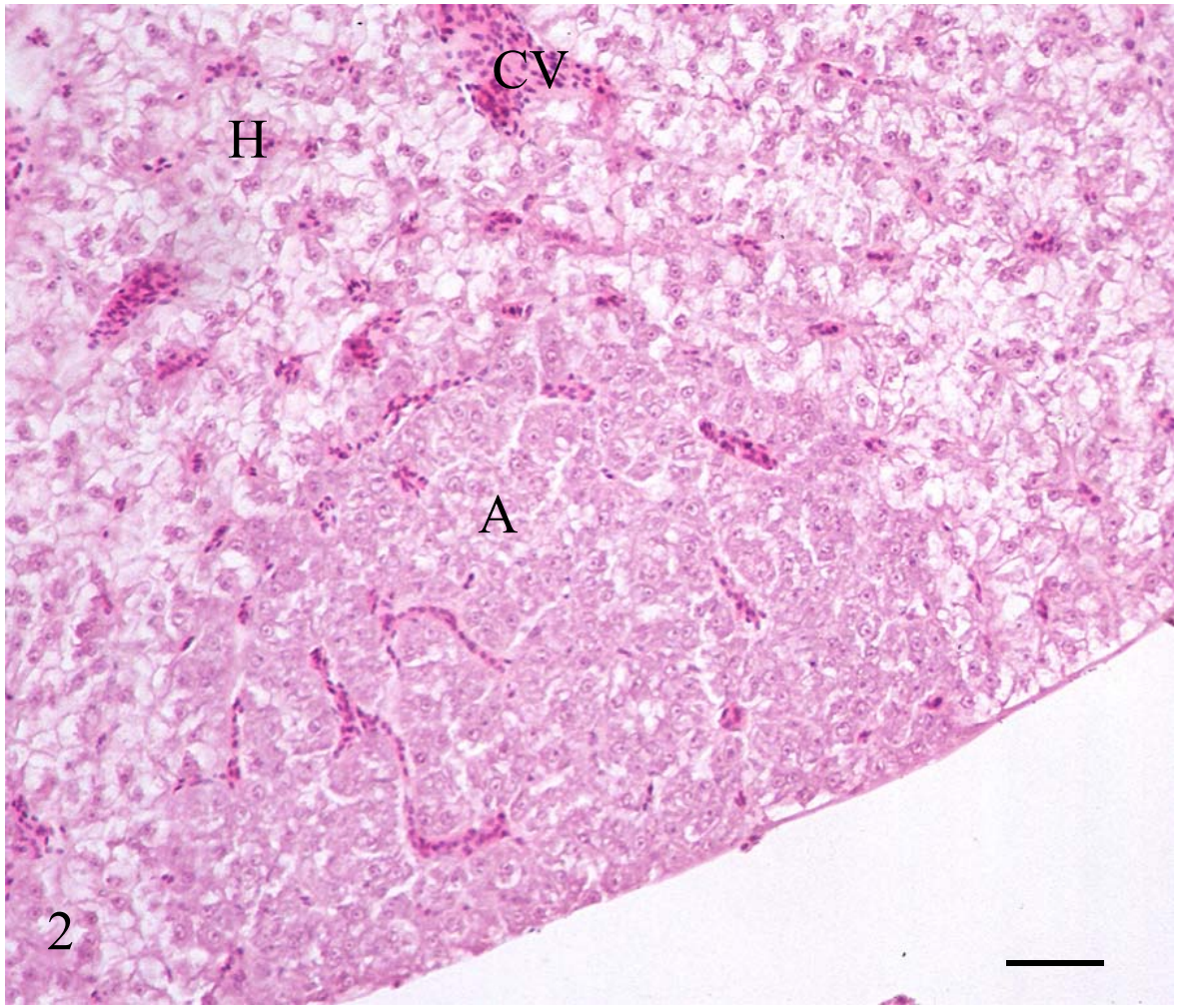


Figure 2. Liver of a one year old medaka exposed to 150 mg/L benzene showing adenoma (A), central vein (CV), and hepatocytes (H). H&E, bar=50 μm.

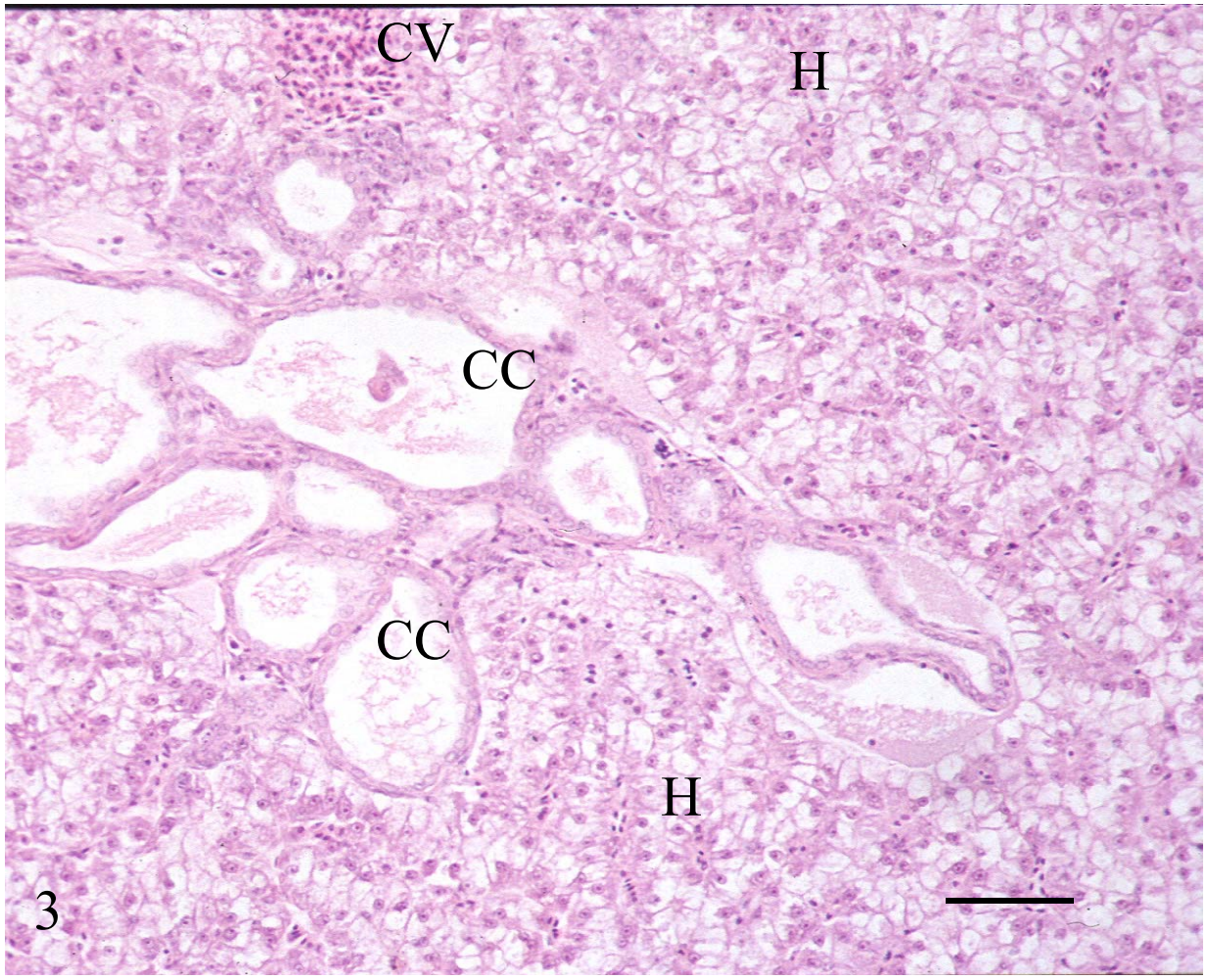


Figure 3. Liver of a one year old medaka exposed to 150 mg/L benzene showing cystic cholangioma (CC), hepatocytes (H), and central vein (CV). H&E, bar=60 μm.

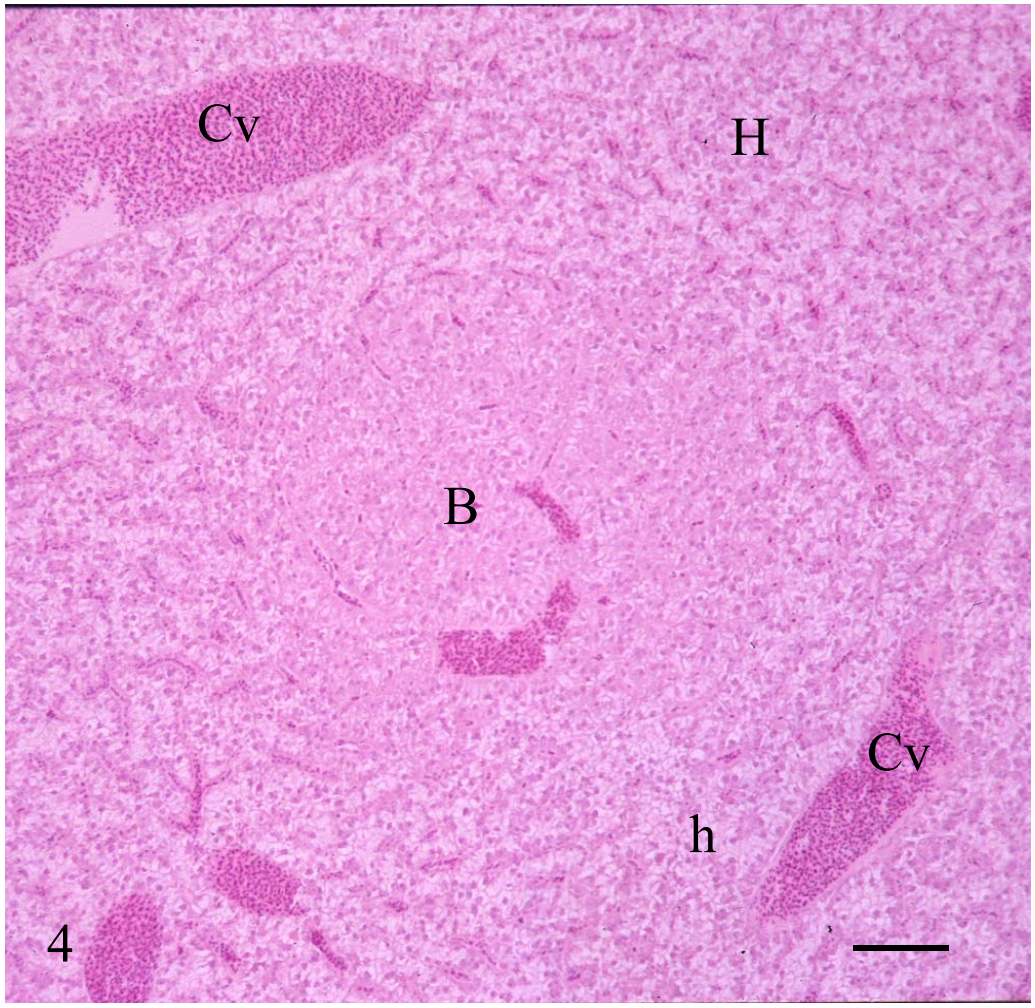


Figure 4. Liver of a one year old medaka exposed to 150 mg/L benzene showing basophilic foci (B), hepatocytes (H), and central vein (CV). H&E, bar=60 μm.

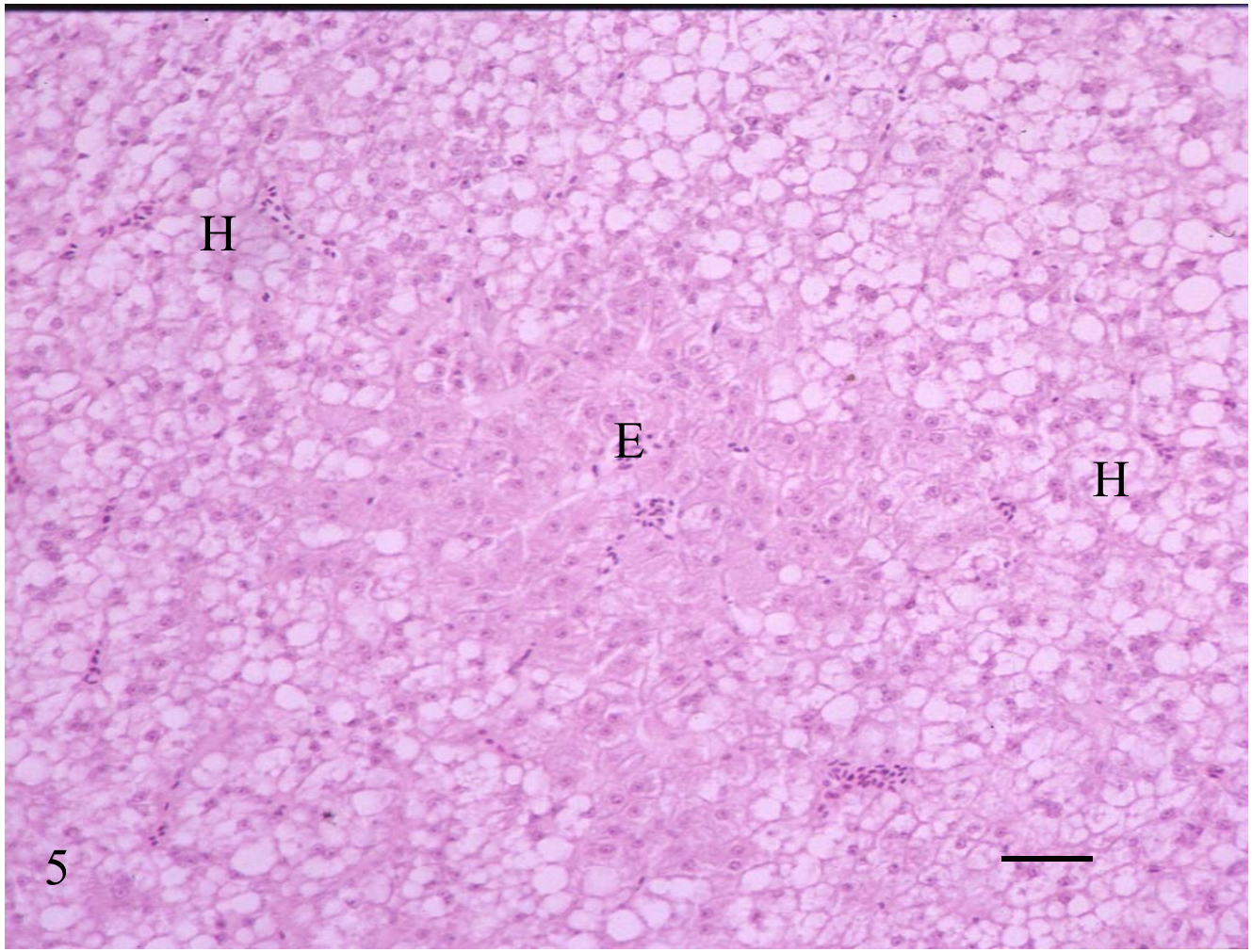


Figure 5. Liver of a one year old medaka exposed to 100 mg/L benzene showing eosinophilic foci (E), and hepatocytes (H). H&E, bar=40 μm.

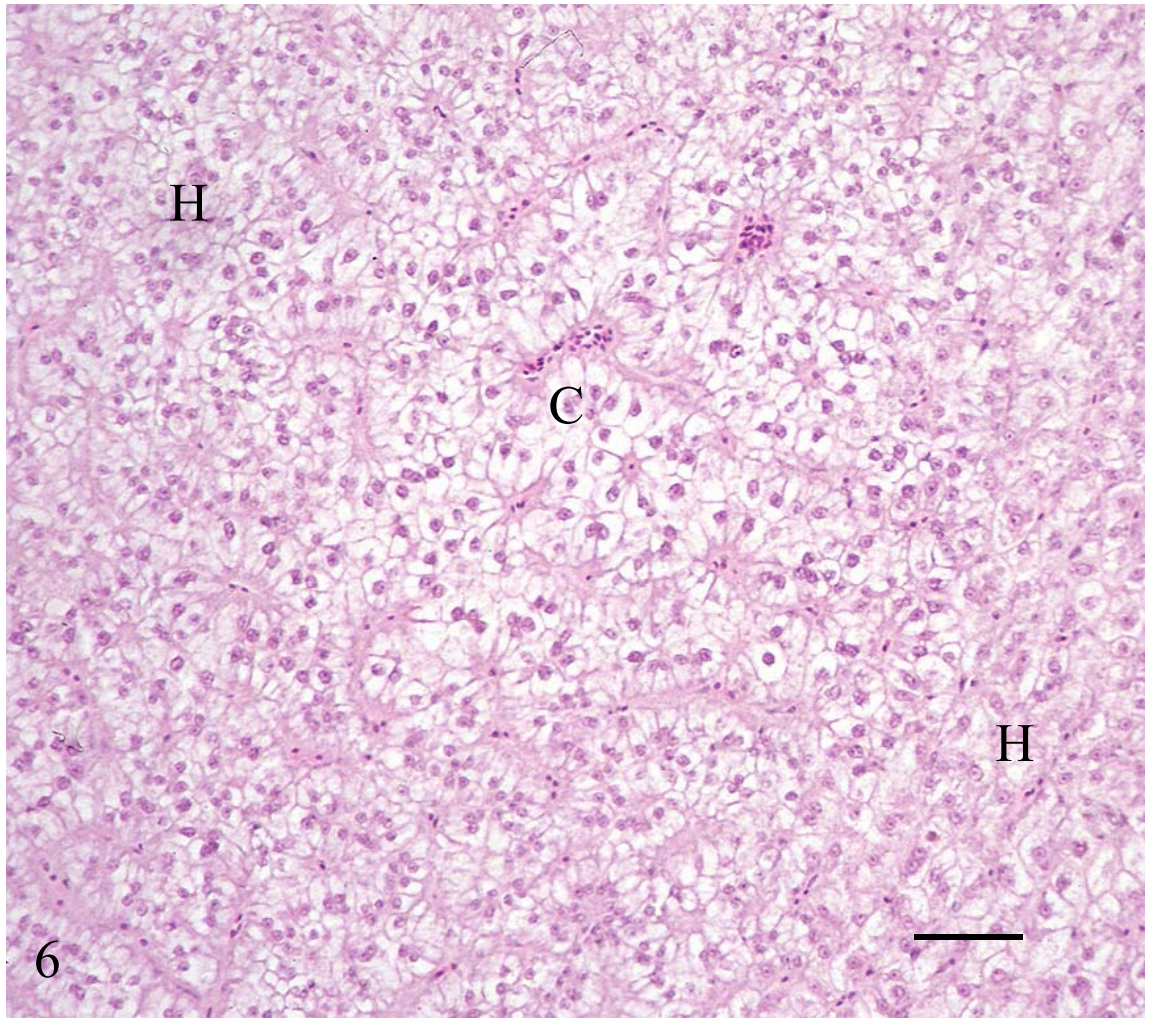


Figure 6. Liver of a one year old medaka exposed to 150 mg/L benzene showing clear cell foci (B),and hepatocytes (H). H&E, bar=60 μm.

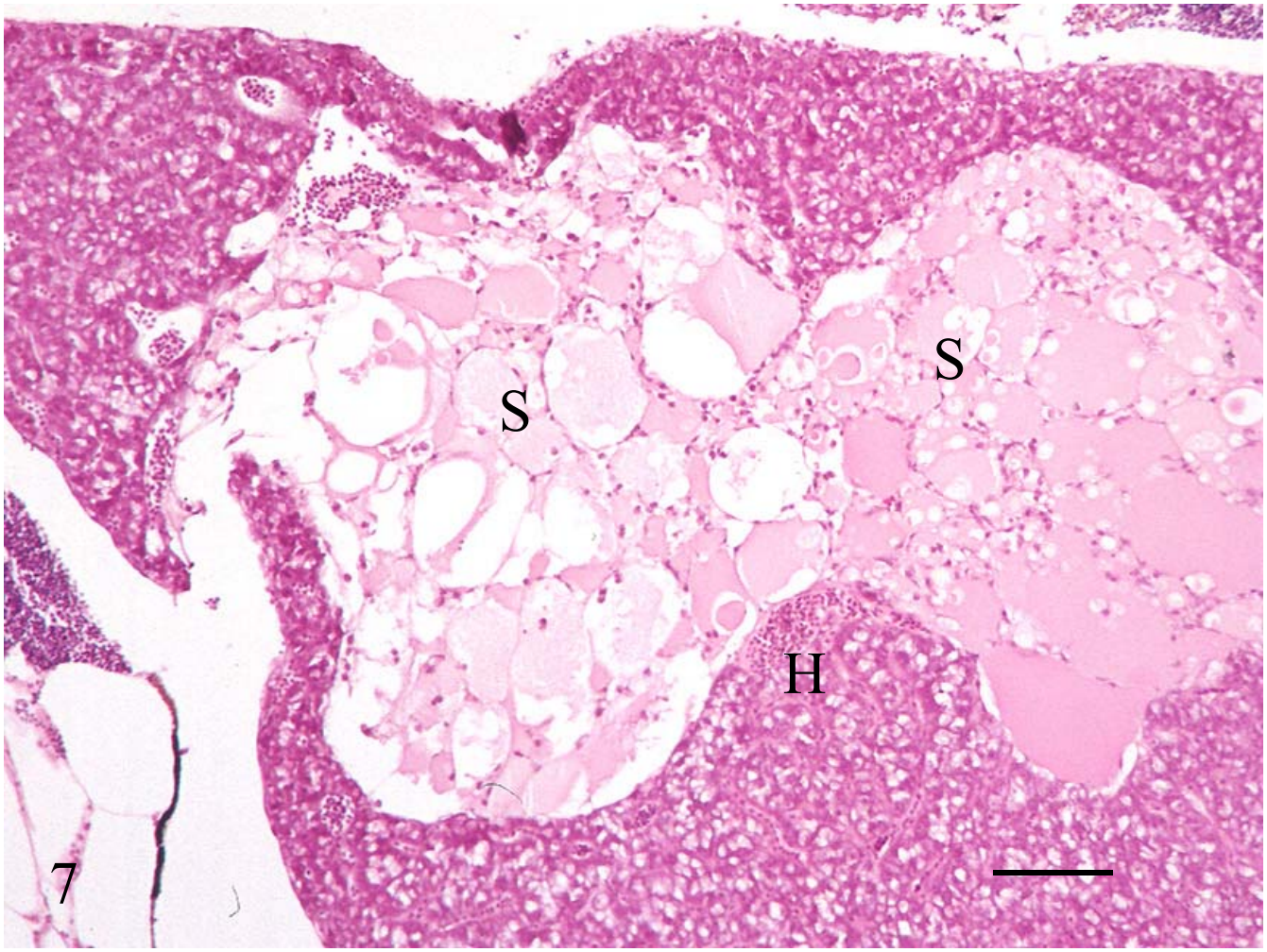


Figure 7. Liver of a one year old medaka exposed to 100 mg/L benzene showing spongiosis hepatis (S) and hepatocytes (H). H&E, bar=60 μm.

The only liver tumors noted in the one-year-old medaka exposed as embryos to benzene were hepatic adenoma and cholangioma. There was an increasing trend (p values $\leq$0.1) in the total tumor incidence in fish surviving after one year in the two high concentration (100 and 150 mg/L benzene) groups. This marginal statistically increased finding also considers that there was one case of hepatic cholangioma in the control group. Based upon the known sensitivity of the medaka to known carcinogens in lifetime (chronic) exposure studies and frequent use in cancer research (Hawkins et al., 1985; Chen et al., 1996), we have attempted an embryo exposure route to determine the sensitivity of the medaka to the carcinogenicity of benzene. Because we were limiting our exposure period to 10 days of embryogenesis and developing new methods, we decided to use high (shock) concentrations of benzene of known embryo lethality to optimize penetration of benzene through the chorion. The hepatocellular adenomas and cholangiomas observed in this study are benign neoplasms. Huff et al. (1989), in evaluating the incidences of selected lesions in two-year gavage studies of benzene in 60 F344/N rats and 60 B6C3F1 mice, reported a non-significant incidence of hepatocellular adenoma similar to our results in vehicle control and all dose groups in both male and female mice. The malignant neoplasms associated with benzene exposure in the previously summarized rodent literature were not observed in our study. Considering the marginal statistical significance of the increased liver tumors and the absence of malignant liver tumors, this medaka embryo exposure methodology does not indicate the known carcinogenic potential of benzene and does not reflect the highly variable carcinogenic response reported in numerous rodent bioassays. The low incidence of non-neoplastic proliferative and degenerative lesions observed in this study are probably a consequence of normal aging. To further test and evaluate the utility of the fish embryo exposure methodology for risk assessment, we are using a similar protocol for the potent animal carcinogen diethylnitrosamine (DEN). We are also considering extending evaluation of our short-term exposure methodology to include the young fish in the age range of several weeks post hatch.

References

Al-Sabti. 1985. Frequency of chromosomal aberrations in the rainbow trout salmo gairdneri rich exposed to fire pollutants. Journal Fish Biology 26: 13-19.

APHA (American Public Health Association), American Water Works Association and Water Environmental Federation. 1992. Standard methods for examination of water and wastewater, 18th edition. APHA, Washington, D.C.

Boorman, G.A., Botts, S., Bunton, T.E. et al. 1997. Diagnostic criteria for degenerative, inflammatory, proliferative nonneoplastic and neoplastic liver lesions in medaka (Oryzias latipes): consensus of a National Toxicology Program Pathology Working Group. Toxicology Pathology 25(2): 202-210.

Brocksen, RW., and Bailey, HC. 1973. Respiratory response of juvenile chinook salmon and striped bass exposed to benzene, a water-soluble component of crude oil. In API Joint Conference on Prevention and Control of Oil Spills. American Petroleum Institute, Washington, DC. p. 783-791.

Chen, HC., Pan, IJ., Tu, WJ., Lin, WH., Hong, CC., and Brittelli, MR. 1996. Neoplastic response in Japanese medaka and Channel catfish exposed to N-methyl-N'Nitro-N-nitrosoguanidine. Toxicologic Pathology 24: 696-706.

Correa, M., and Garcia, H. 1990. Physiological responses of juvenile white mullet*, Mugil curema,* exposed to benzene. Bulletin of Environmental Contamination and Toxicology 44: 428-434.

Cronkite, EP. 1986. Benzene hematotoxicity and leukemogenesis. Blood Cells 12: 129-137.

Cronkite, EP., Bullis, JE., Inoue, T. et al., 1984. Benzene inhalation produces leukemia in mice. Toxicology and Applied Pharmacology 75: 358-361.

Cronkite, EP., Drew, RT., Inoue, T. et al., 1985. Benzene hematotoxicity and leukemogeneisis. American Journal of Industrial Medicine 7: 447-456.

Cronkite, EP., Drew, RT., Inoue, T. et al., 1989. Hematotoxicity and carcinogenicity of inhaled benzene. Environmental Health Perspectives 82: 97-108.

Farris, GM., Everitt, JI., Irons, RD., et al. 1993. Carcinogenicity of inhaled benzene in CBA mice. Fundamental and Applied Toxicology 20: 503-507.

Hartley, WR., Thiyagarajah, A., and Mizell, M. 1995. Ectopic hermaphroditic gonad in the cranial cavity of Japanese medaka. Journal of Aquatic Animal Health 7: 172-177.

Hawkins, WE., Overstreet, RM., Fournie, JW., et al., 1985. Development of aquarium fish models for environmental carcinogens: Tumor induction in seven species. Journal of Applied Toxicology 5: 261- 264.

Huff, JE., Haseman, JK., DeMarini, DM., Eustis, S., Maronpot, RR., Peters, AC., Persing, RL., Chrisp, CE., and Jacobs, AC. 1989. Multiple-site carcinogenicity of benzene in Fischer 344 rats and B6CF1 mice. Environmental Health Perspectives 82: 125-163.

Hunting, KL., Longbottom, H., Kalavar, SS., et al. 1995. Haematopoietic cancer mortality among vehicle mechanics. Occupational and Environmental Medicine 52: 673-678.

Hurley, JF., Cherrie, JW., and Maclaren, W. 1991. Exposure to benzene and mortality from leukaemia: Results from coke oven and other coal product workers. British Journal of Industrial Medicine 48: 502-503.

Kirchen, R., and West, W. 1979. The Japanese Medaka, Its Care and Development, Burlingame, NC: Carolina Biological Supply Co.

Maltoni. C., Cotti, G., Valgimigli, L., et al. 1982a. Hepatocarcinomas in Sprague-Dawley rats following exposure to benzene by inhalation: First experimental demonstration Med Lav 4: 446-450.

Maltoni. C., Cotti, G., Valgimigli, L., et al. 1982b. Zymbal gland carcinomas in rats following exposure to benzene by inhalation. American Journal of Industrial Medicine 3: 11-16.

Maltoni. C., Conti, B., and Cotti, G. 1983. Benzene: A multiple carcinogen: Results of a long-term bioassays performed at the Bologna Institute of Oncology. American Journal of Industrial Medicine 4: 589-630.

Maltoni. C., Conti, B., Cotti, G., et al. 1985. Experimental studies on benzene carcinogenicity at the Bologna Institute of Oncology: Current results and ongoing research. American Journal of Industrial Medicine 7: 415-446.

Maltoni. C., Ciliberti, A., Cotti, G., et al. 1989. Benzene, an experimental multipotential carcinogen: Results of the long-term bioassays performed at the Bologna Institute of Oncology. Environmental Health Perspectives 82: 109-124.

Meyerhoff, RD. 1975. Acute toxicity of benzene, a component of crude oil, to juvenile striped bass, *Morone saxatilis*. Journal of Fisheries Research Board of Canada 32: 1864-1866.

Purcell, WP. 1978. Benzene. In: Grayson, M., and Eckroth, D., (eds.) Kirk-Othmer Encyclopedia of Chemical Technology, 3rd edition, Vol. 3. New York: John Wiley and Sons, p. 744-771.

Rugh, R. 1962. Experimental Embryology Techniques and Procedures, third edition, Burgess Publishing Co., Minneapolis.

Snyder, CA., Goldstein, BD., Sellakumar, AR. 1978. Hematotoxicity of inhaled benzene to Sprague-Dawley rats and AKR mice at 300 ppm. Journal of Toxicology and Environmental Health 4: 605-618.

Snyder, CA., Goldstein, BD., Sellakumar, AR. 1980. The inhalation toxicology of benzene: Incidence of hematopoietic neoplasms and hematotoxicity in AKR/J and C57BL/6J mice. Toxicology and Applied Pharmacology 54: 323-331.

Snyder, CA., Goldstein, BD., Sellakumar, AR. 1984. Evidence for hematotoxicity and tumorigenesis in rats exposed to 100 ppm benzene. American Journal of Industrial Medicine 5: 429-434.

U.S. EPA. 1999. Integrated Risk Information System (IRIS). Online. National Center for Environmental Assessment, Washington, DC.

PROGRESS REPORT: CHRONIC TOXICITY OF THE DISINFECTION BY-PRODUCTS (DBPS) CHLOROFORM AND BROMOFORM IN THE JAPANESE MEDAKA (*ORYZIAS LATIPES*)

William R. Hartley, Sc.D. and A. Thiyagarajah, Ph.D.

Tulane University
School of Public Health and Tropical Medicine
New Orleans, LA

September 1999

Abstract

The trihalomethanes (THMs), chloroform, bromoform, bromodichloromethane (BDCM) and chlorodibromomethane (CDBM), are present in drinking water in relatively large quantities, compared with other DBPs. Because epidemiologic studies suggest the potential for cancer and developmental effects in humans from DBP exposures, the THMs were selected for evaluation. Japanese medaka embryos were exposed for the first 10 days of embryological development to chloroform (25, 50, or 100 ppm) or bromoform (5, 10, 25, or 50 ppm). Two control groups were used, one in an embryo-rearing solution (ERS) and one in ERS/dimethyl sulfoxide (DMSO). At 6 and 12 months, the fish were evaluated for chronic effects including cancer. Preliminary results are given in this progress report.

For chloroform, a low incidence of neoplasms in the liver was observed. After 12 months, cholangioma was found in two fish at 25 mg/L and one fish at 100 mg/L, and lipoma was found in one fish at 50 mg/L. Non-neoplastic proliferative lesions observed were clear cell foci, vacuolated cell foci, and altered hepatocellular altered foci. Excessive accumulation of substances such as glycogen or lipid occurred in large numbers of fish after 6 months; the incidence decreased in fish examined after 12 months. In addition, hepatic cysts and spongiosis hepatis were found in high incidence. In the kidney, the incidence of cystic tubules in the control and exposed fish were similar. Other tubular lesions observed were accumulation of hyaline globules in the tubular epithelium and a lower incidence of vacuoles in the tubular epithelium and tubular necrosis. In the thyroid, there was increased infollicular hyperplasia at 12 months. Delayed sexual maturation was observed at all concentrations including controls. Immature testis were found only in fish examined after 6 months; none of the 12-month-old fish had immature testis except one control fish. The incidence of immature testis in fish examined after 6 months was higher than controls. In female fish, immature ovaries occurred in all groups including controls, with the incidence higher in the exposed fish.

For bromoform, the incidence of liver lesions was similar to that observed in fish exposed to chloroform. Incidence of hepatic cysts was high compared to controls. One fish exposed at 50 mg/L for six months had cholangioma. In the kidney, a high incidence of cystic tubules occurred in all groups of fish including controls. In the thyroid, a high incidence of follicular hyperplasia occurred in 12-month-old fish including controls compared to 6-month-old fish. A high incidence of immature testis and ovaries occurred in 6-month-old fish exposed at 10 mg/L.

At this time (September 1999), this progress report provides the largely complete laboratory results on chloroform and significant information on bromoform. Statistical analyses have not yet been performed on these data. Completion of the 12-month data sets for bromoform at 10 mg/L and DMSO are in progress. The 50 mg/L bromoform group was terminated prior to 12 months because of high mortality. The 5 and 25 mg/L bromoform groups (6 and 12 months) are currently under histological evaluation. Similar chronic experiments for BDCM are in progress. This progress report does not include an update on cardiac lesions (dilated atrium and tubular heart) as special sectioning of the tissues is being performed to more clearly understand the etiology of the high incidence of these lesions.

Introduction

Water treatment systems use chemical disinfectants (chlorine, ozone, and /or chloramination) to eliminate infectious agents from drinking water. These chemicals produce a variety of by-products often called disinfection by-products (DBPs). The common ones are trihalomethanes, haloacetic acids, haloacetonitriles, haloketones, and aldehydes. We selected the four trihalomethanes, chloroform, bromoform, bromodichloromethane (BDCM) and chlorodibromomethane(CDBM), because of the toxicological concerns related to these chemicals. The concentrations and the ratios of these four trihalomethanes vary with source water quality, treatment system designs, and seasonal water quality parameters. Japanese medaka (*Oryzias latipes*) embryos were exposed for the first 10 days of embryological development to chloroform and bromoform separately. At 6 and 12 months, they were evaluated for chronic effects including cancer.

This progress report provides the largely complete results on chloroform and significant information on bromoform (6 and 12 months). Completion of the 12-month data sets for the bromoform 10 mg/L concentration group and dimethyl sulfoxide (DMSO) are in progress. The 50 mg/L bromoform concentration group was terminated prior to 12 months because of high mortality. The 5 mg/L and 25 mg/L bromoform concentration groups (6 and 12 months) are currently under histological evaluation. Similar chronic experiments for BDCM are in progress. Both the bromoform and chloroform experiments were designed using the same control groups of embryo-rearing solution (ERS) and ERS/DMSO.

In addition, progress report does not include an update on cardiac lesions (dilated atrium and tubular heart) in chloroform- and bromoform-exposed fish. We are currently doing special sectioning of the tissues to more clearly understand the etiology of the high incidence of these lesions.

Methods

Medaka Culture

Tulane University has an established broodstock medaka fish culture. Primary culture conditions for the broodstock fish follow the general guidance provided by Kirchen and West (1979) and others. Medaka are kept in 20-gallon plexiglass aquaria with recirculating filters. The photoperiod is 14 hours light/10 hours dark, and the temperature is maintained between 21 to 25˚C. Fertilized eggs are collected 1 to 2 hours after the beginning of the 14-hour light cycle. Breeding fish are maintained on a high protein commercial flake food and weekly tonics of live or frozen brine shrimp. Water quality is monitored for pH (adjusted to 7.0 to 7.5), ammonia (<0.5 mg/L), and dissolved oxygen (7-8 mg/L).

Exposure Methodology for Chronic Toxicity Studies

Eggs were removed from fish by using microdissection forceps. Eggs were immediately placed into a sterile embryo growth medium (Kirchen and West 1979; Rugh, 1962), dechorionated by mechanical abrasion, and then pooled together in a glass bowl. The viability of the embryos was carefully evaluated, and the unfertile eggs were discarded. Randomly selected embryos at the early blastula stage were exposed to three nominal concentrations of chloroform (25, 50, or 100 ppm) or four concentrations of bromoform (5, 10, 25, or 50 ppm) in groups for 10 days. Two control groups (spring water control and a carrier solvent-DMSO control) were used in this study. The exposure solutions were renewed (80%) daily throughout the exposure period. At the end of a 10-day exposure, embryos were washed thoroughly and then transferred to clean embryo-rearing solution until they hatch. The fry were transferred to grow-out tanks (10-gallon recirculating aquaria). Fish were fed once daily with commercial flake food. Water quality was monitored and maintained as described in medaka culture section of this report.

Sampling and Processing

Sixty fish were randomly sampled at the end of a 6 month grow-out period and the survivors were sampled at the end of 12 months. Fish were killed by an overdose of MS222, and then the abdomen was cut opened. Fish were fixed as whole in 10% neutral buffered formalin. Fish were then decalcified in a formic acid:sodium citrate mixture and then processed for histological evaluation using standard histological procedures. Sagittal sections were cut at 5 μm thick and stained with hematoxylin and eosin. Two left paramedian sections and two midsagittal sections were cut and stained. The right half of the fish was archived according to a method modified from Wolfe (1994). Evaluations of these samples are still ongoing. The available data are presented in this report.

Results and Discussion

Chloroform Studies

Liver

Liver lesions were grouped into 1) neoplasms, 2) nonneoplastic proliferative lesions, and 3) degenerative lesions using the criteria described by Boorman et al. (1998). The incidence of liver lesions in medaka examined after 6 and 12 months are in Table 1. In all tables of data presented, differences in the Total N and the number examined for each organ system result from the organ not being present in the fish sections (slides) prepared. In such cases, the fish are being resectioned to obtain the desired tissue. Figure 1 illustrates the morphology of a control medaka liver.

Table 1

Incidence of Liver Lesions in 6- and 12-Month-Old Fish
Exposed to Chloroform as Embryos

Lesions	ERS		DMSO	25 ppm		50 ppm		100 ppm	
	6M	12M	6M	6M	12M	6M	12M	6M	12M
Total N	60	60	60	60	77	60	62	60	32
Number examined	52	42	47	59	73	60	62	58	27
No lesions	45	26	39	32	37	29	41	41	18
Degenerative Lesions									
Excessive accumulation of substances	2	0	1	23	0	12	8	10	2
Fatty change	1	3	2	0	3	4	4	1	0
Hepatic cysts	5	4	5	0	7	11	6	3	4
Spongiosis hepatis	0	4	1	0	13	2	0	0	3
Granuloma	0	0	0	1	14	0	4	0	0
Necrosis	0	4	0	1	2	2	2	1	1
Non-neoplastic Proliferative Lesions									
Bile duct hyperplasia	0	0	0	0	0	0	1	0	0
Biliary cyst	0	0	0	1	0	2	0	0	0
Altered foci	0	0	0	0	0	1	0	0	0
Clear cell foci	0	1	0	2	1	0	0	1	0
Vacuolated cell foci	0	2	0	0	3	0	0	0	0
Neoplasms									
Cholangioma	0	0	0	0	2	0	0	0	1
Lipoma	0	0	0	0	0	0	1	0	0
Miscellaneous									
Ectopic gonad	0	0	0	0	0	0	0	1	0

ERS = embryo-rearing solution
DMSO = dimethyl sulfoxide
M = months

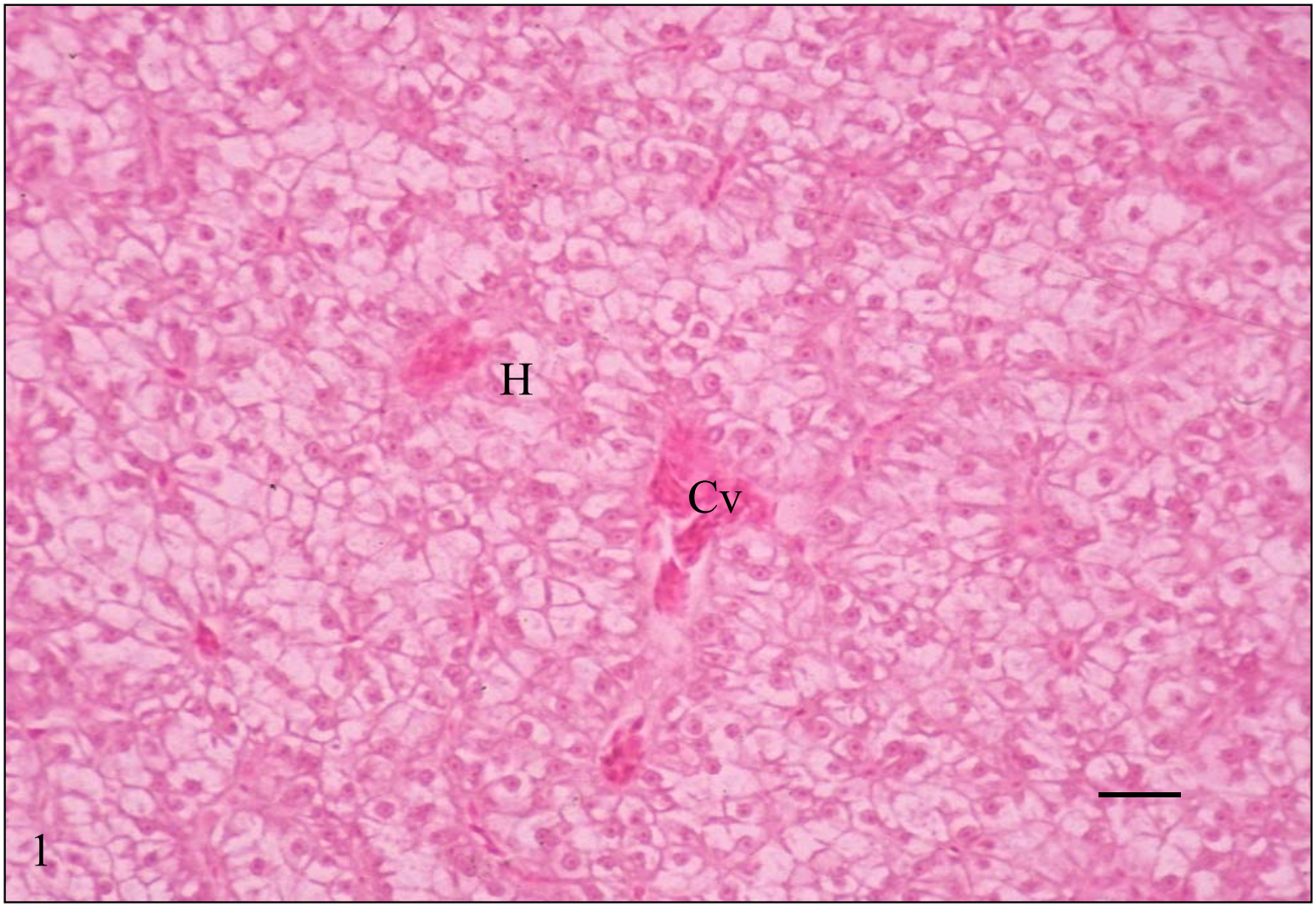


Figure 1. Histomorphology of a control medaka liver. Hepatocytes (H). H&E, bar=20 μm.

A low incidence of neoplasms was observed in this study (Table 1). The types of neoplasms observed were cholangioma and lipoma. Cholangioma (Figure 2) was found in two fish exposed to 25 ppm choloroform and one fish exposed to 100 ppm chloroform and examined after 12 months. Lipoma (Figure 3) was found in one fish exposed to 50 ppm and examined after 12 months. These two tumors are considered benign. Non-neoplastic proliferative lesions observed were clear cell foci, (Figure 4), vacuolated cell foci, and altered hepatocellular altered foci.

Excessive accumulation of substances (Figure 5) such as glycogen or lipid occurred in large numbers of fish examined after 6 months, but the incidence decreased in fish examined after 12 months. There was excessive accumulation of substances in chloroform-exposed fish. A very low incidence was observed in ERS control or DMSO control fish. The nature of the substances was not investigated. Because the hepatocyte cytoplasm appeared clear or vacuolated, it was assumed these substances were glycogen or lipid, which would have dissolved during the processing of tissues. There were ruptures of hepatocytic plasma membrane and cell death as a consequence of excessive accumulation of these substances.

Other degenerative lesions observed were fatty change and necrosis. In addition, hepatic cysts (Figure 6) and spongiosis hepatis (Figure 7) were also found in high incidence. Hepatic cysts were single cysts lined by thin processes of endothelial cells and often contained eosinophilic material. Spongiosis hepatis was characterized by multilocular cystic structures lined by thin processes of endothelial cells and containing eosinophilic material. Hepatic cysts appeared to be precursor lesions for spongiosis hepatis, but the relationship was not clear.

There was a single case of ectopic (found in an abnormal location) ovarian follicles in the liver (Figure 8) in a female fish exposed to 100 ppm chloroform and examined after 6 months. Ectopic ovarian follicles consisted previtellogenic follicles only. A mature ovary was found in the normal location of this fish.

Kidney

Kidney lesions were grouped into 1) tubular lesions, 2) glomerular lesions, and 3) interstitial tissue lesions. Tubular lesions were characterized by either unilocular or multilocular, occasionally papillary, cystic tubules (Figure 9), where the lumen of the tubules were dilated to varying degrees. Sometimes, the tubular lumen contained proteinaceous material (hyaline cast). The incidence of cystic tubules in control fish was similar to that of exposed fish. Therefore, we believe this lesion may not be directly associated with toxic insult, but probably the result of obstructions in the larger tubules or collecting ducts. Other tubular lesions observed were accumulation of hyaline globules (Figure 10) in the tubular epithelium and a lower incidence of vacuoles (Figure 11) in the tubular epithelium and tubular necrosis. Thyroidization of tubules was not observed in any of the chloroform-exposed fish.

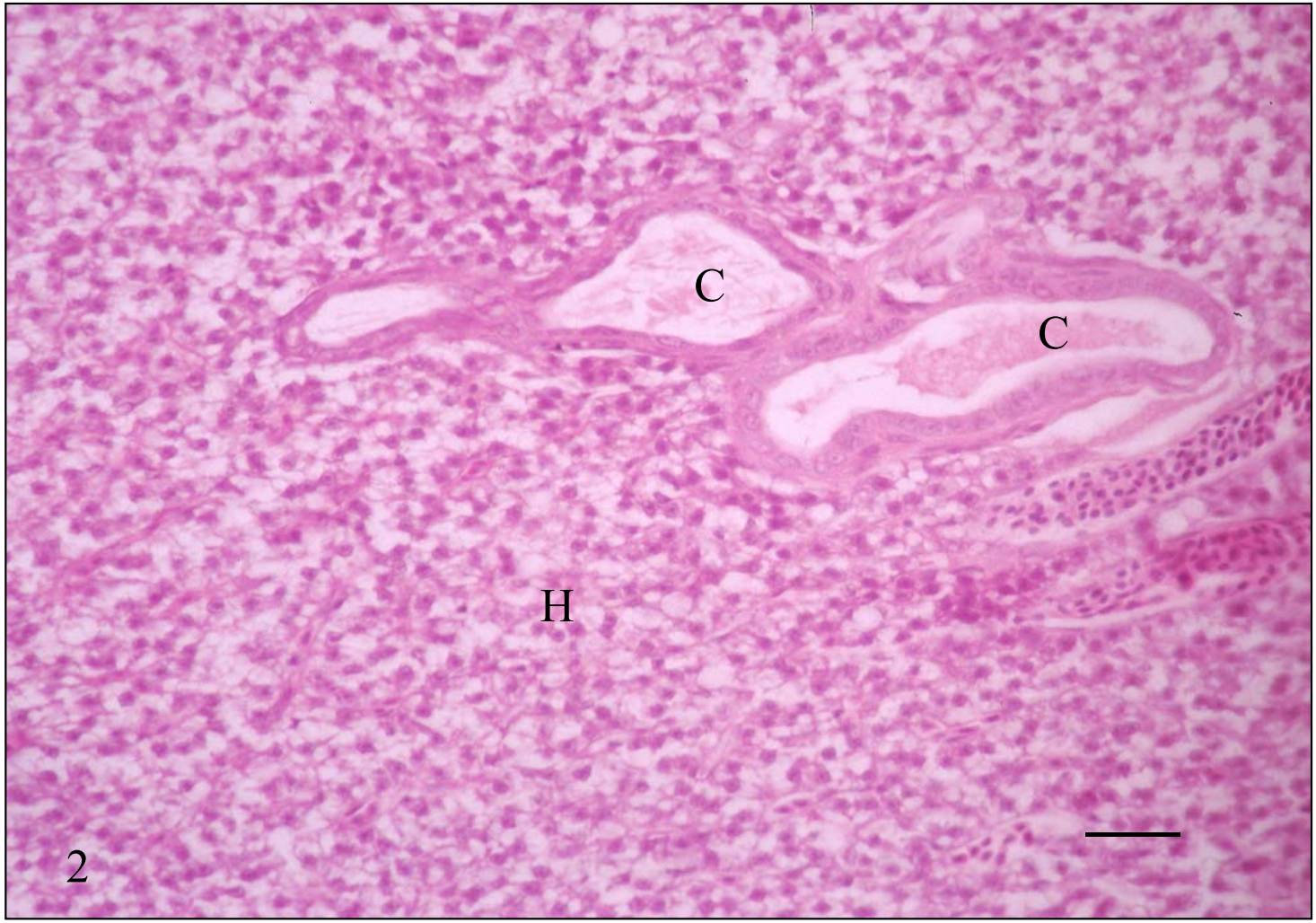


Figure 2. Cholangioma (C) in a medka exposed as embryos to 25 ppm chloroform for 10 days, and evaluated after 12 months. Hepatocytes (H). H&E, bar=20 μm.

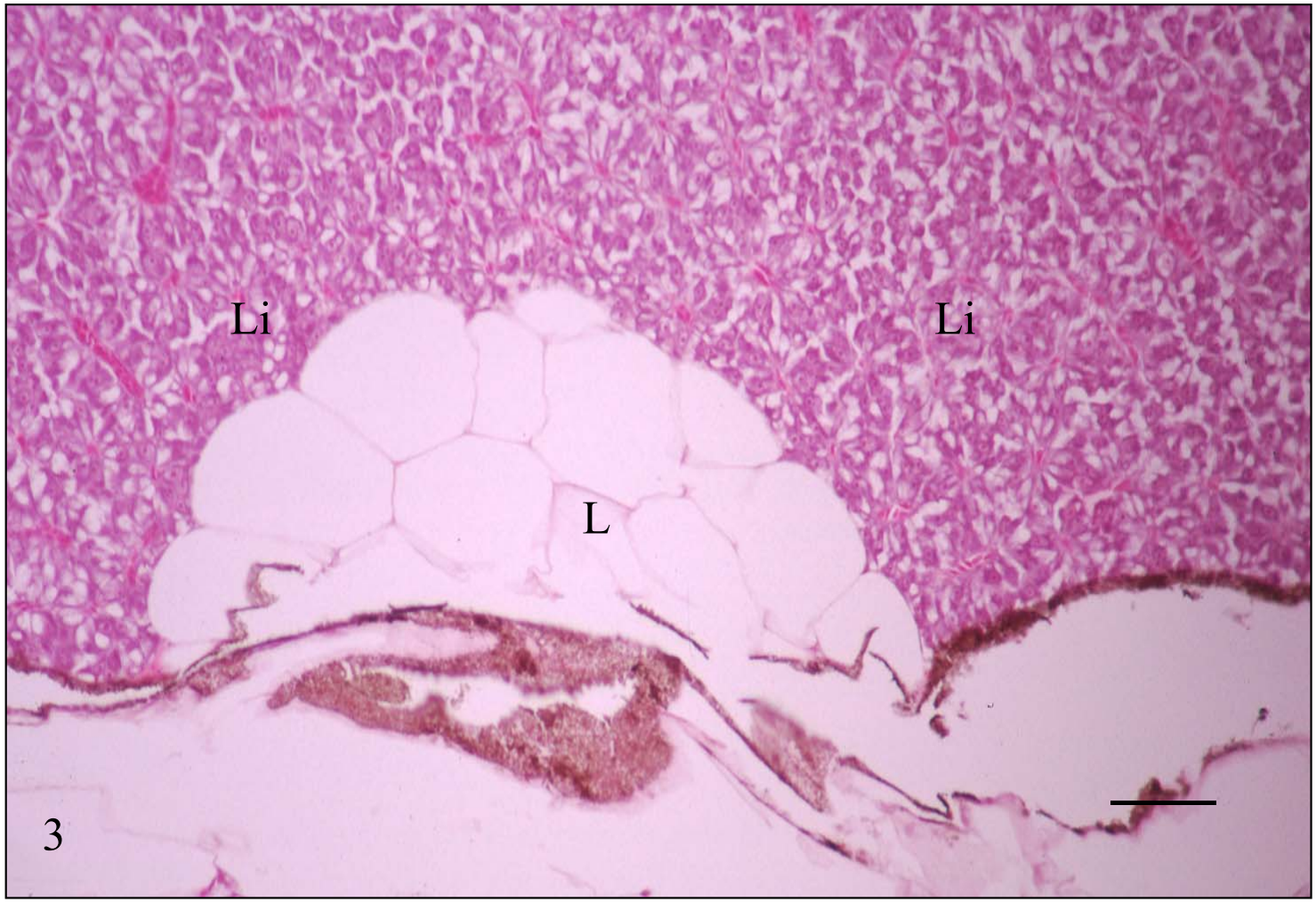


Figure 3. Lipoma (L) in a medaka liver (Li) exposed as embryos to 50 ppm chloroform for 10 days and evaluated after 12 months. H&E, bar=25 μm.

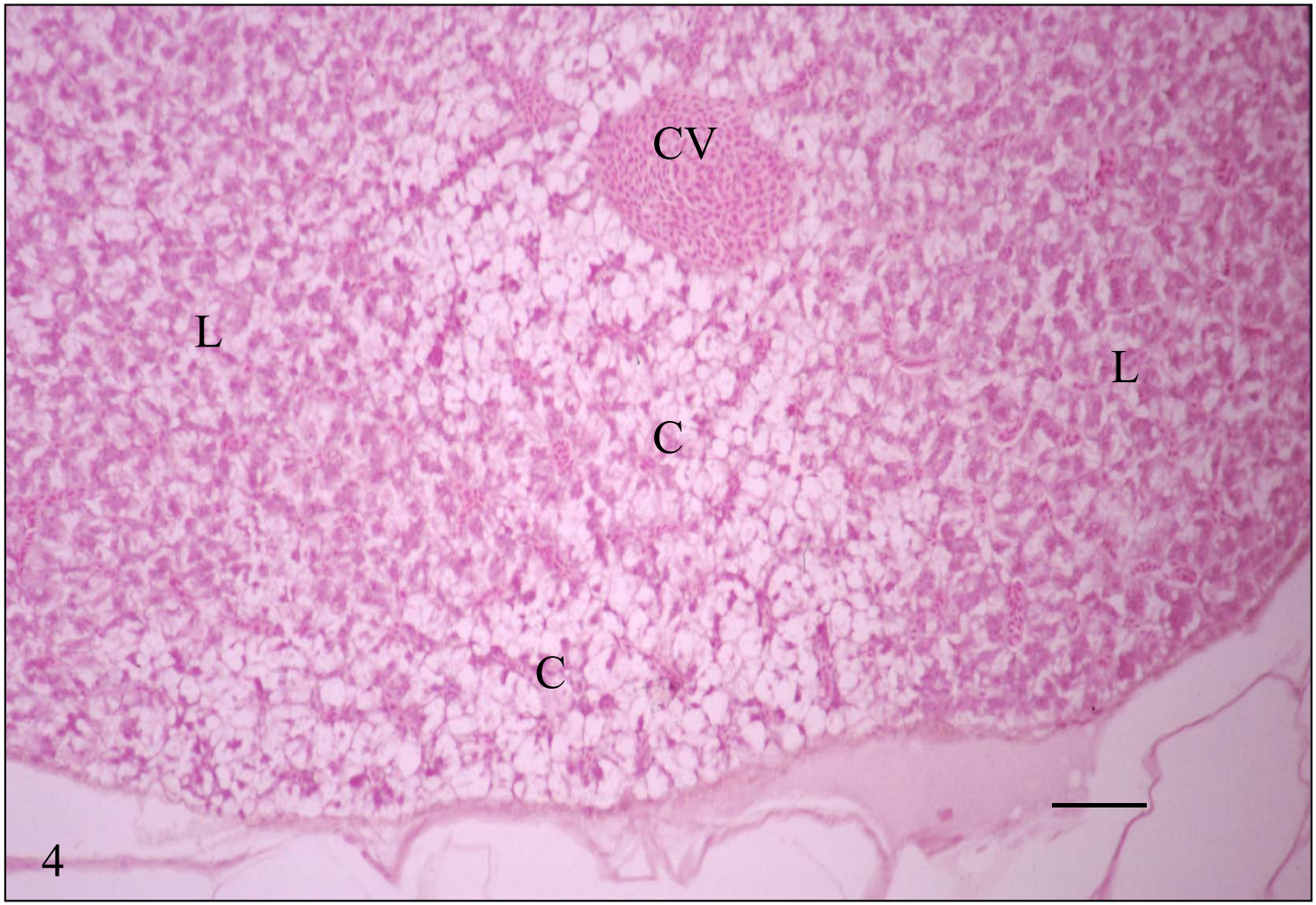


Figure 4. Clear cell foci (C) in a medaka liver (L) exposed as embryos to 10 ppm bromoform for 10 days and evaluated after 12 months. H&E, bar=20 μm.

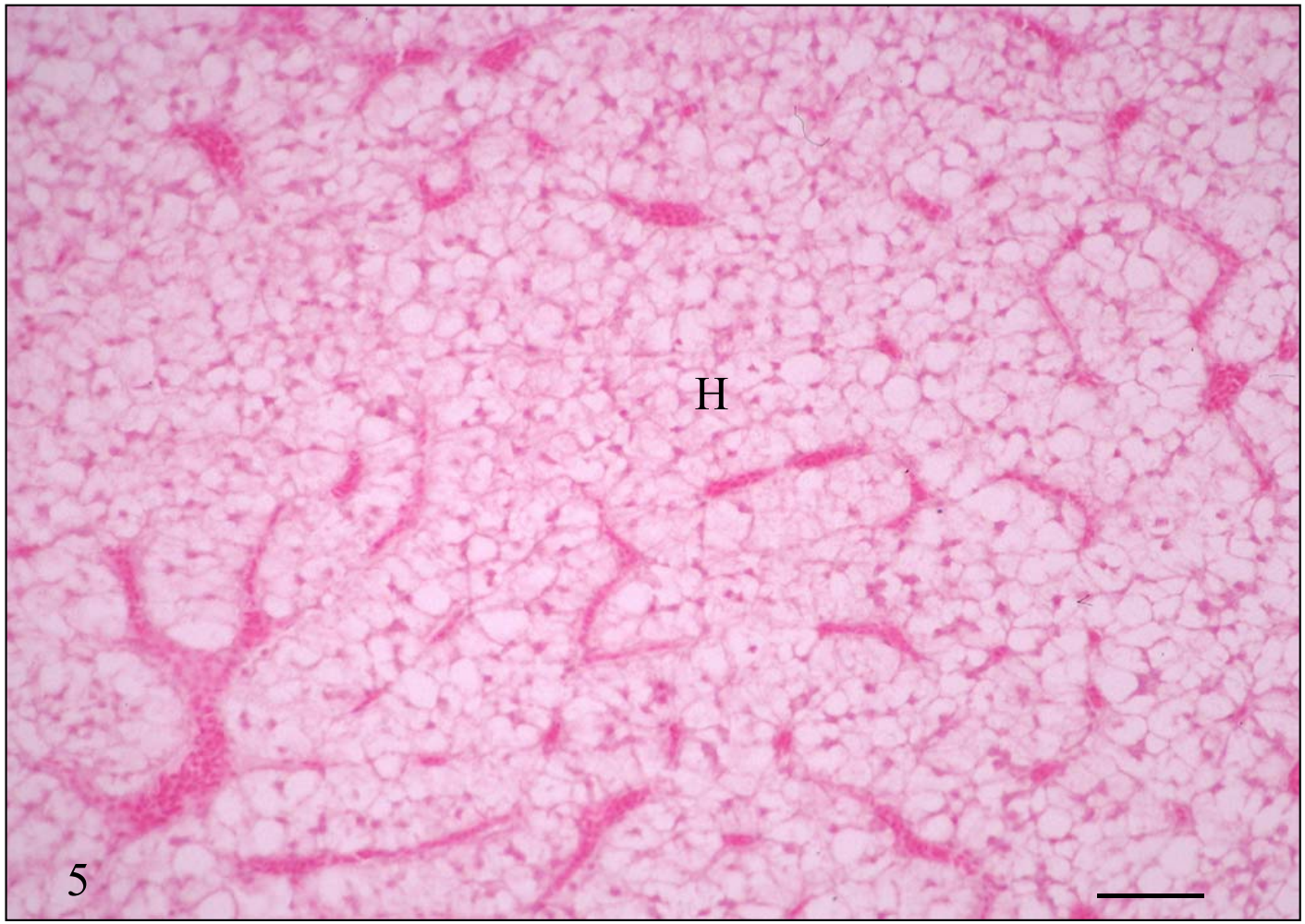


Figure 5. Excessive accumulation of substances in the hepatocyte (H) cytoplasm of a medaka exposed as embryos to 50 ppm bromoform for 10 days and evaluated after 6 months. Note the clear cytoplasm and compressed or necrotic nuclei. H&E, bar=25 μm.

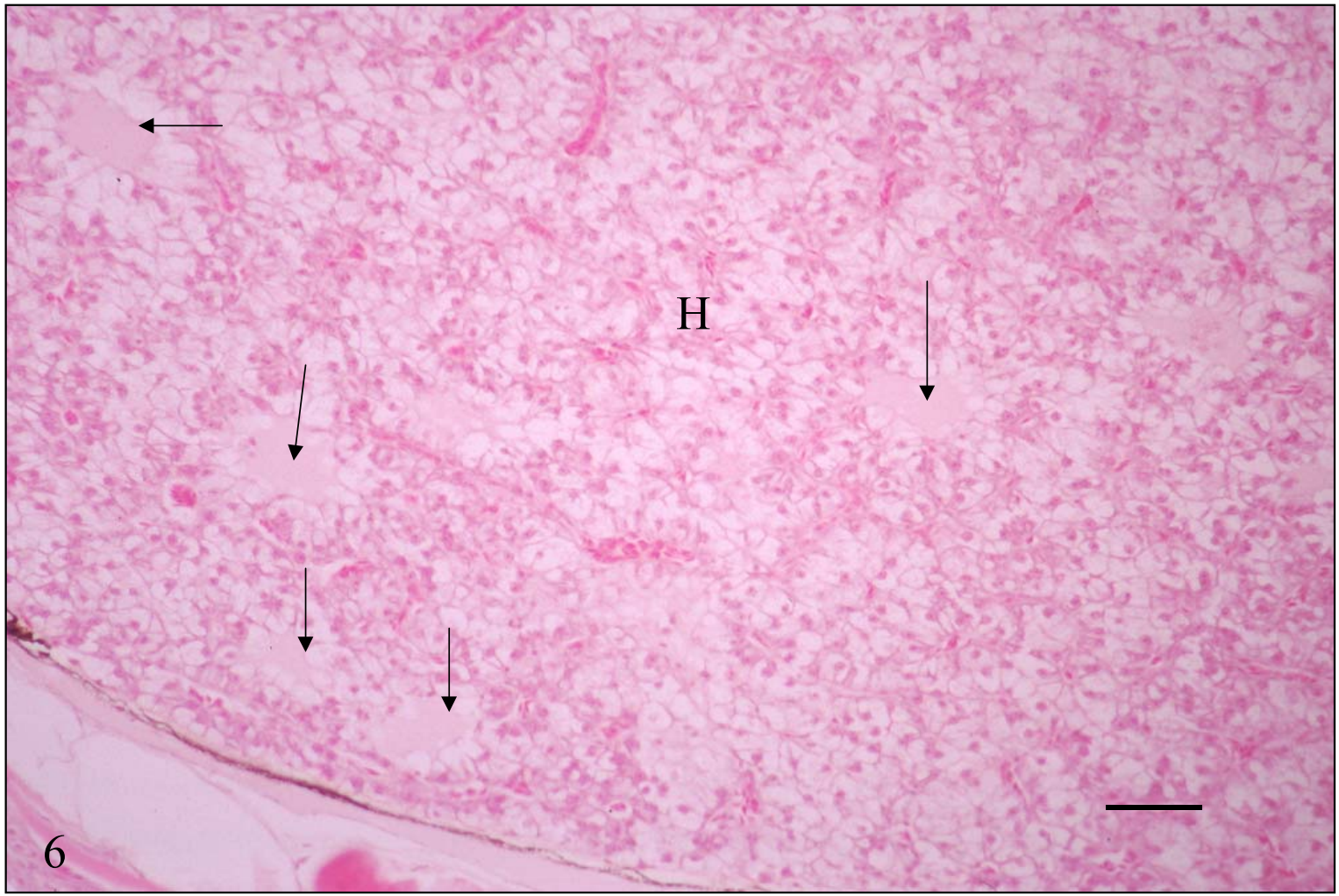


Figure 6. Hepatic cysts (arrows) in a medaka liver (H) exposed to 50 ppm bromoform for 10 days and evaluated after 6 months. H&E, bar=20 μm.

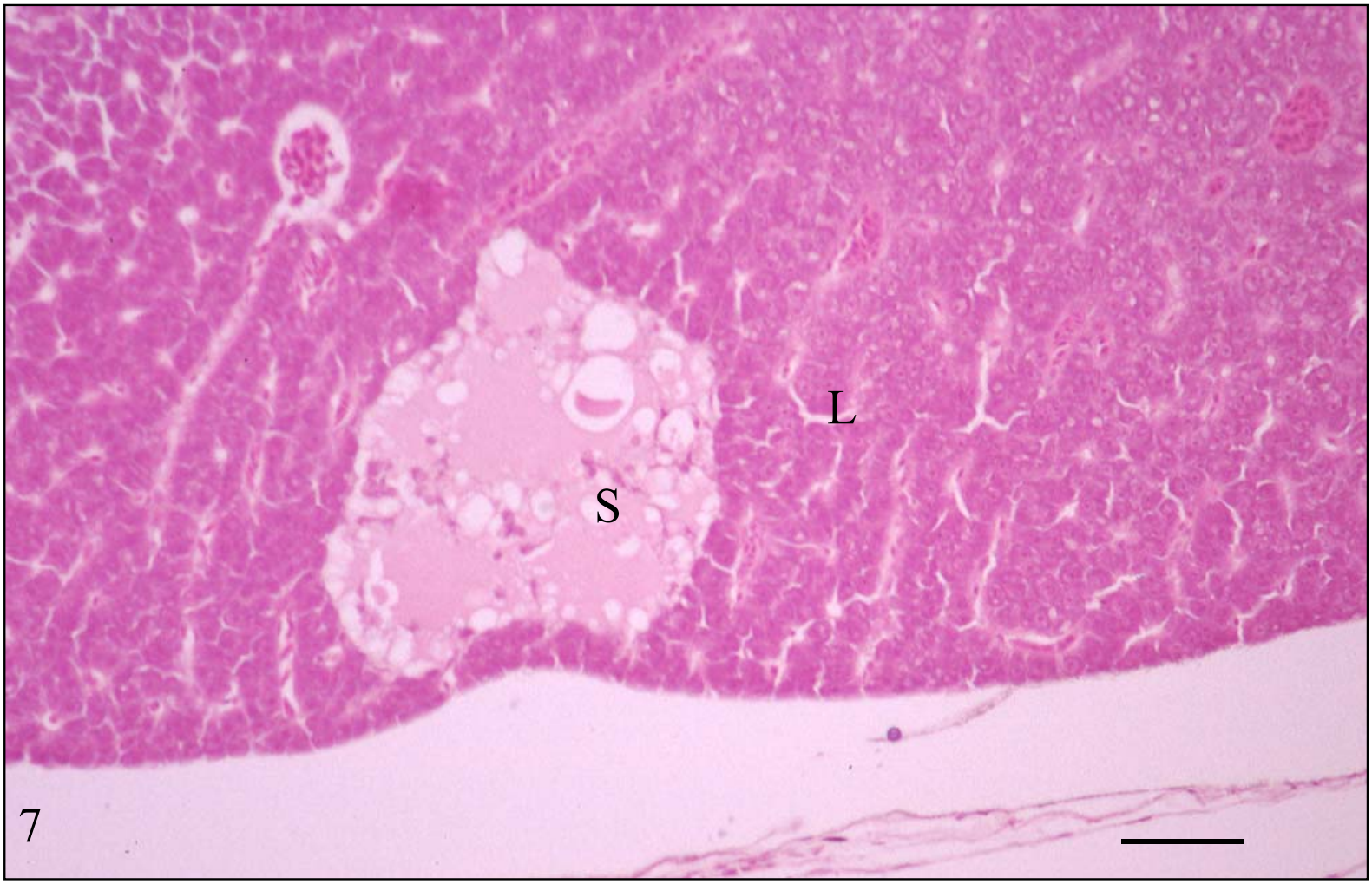


Figure 7. Spongiosis hepatis (S) in a control medaka liver (L) evaluated after 12 months. H&E, bar=30 μm.

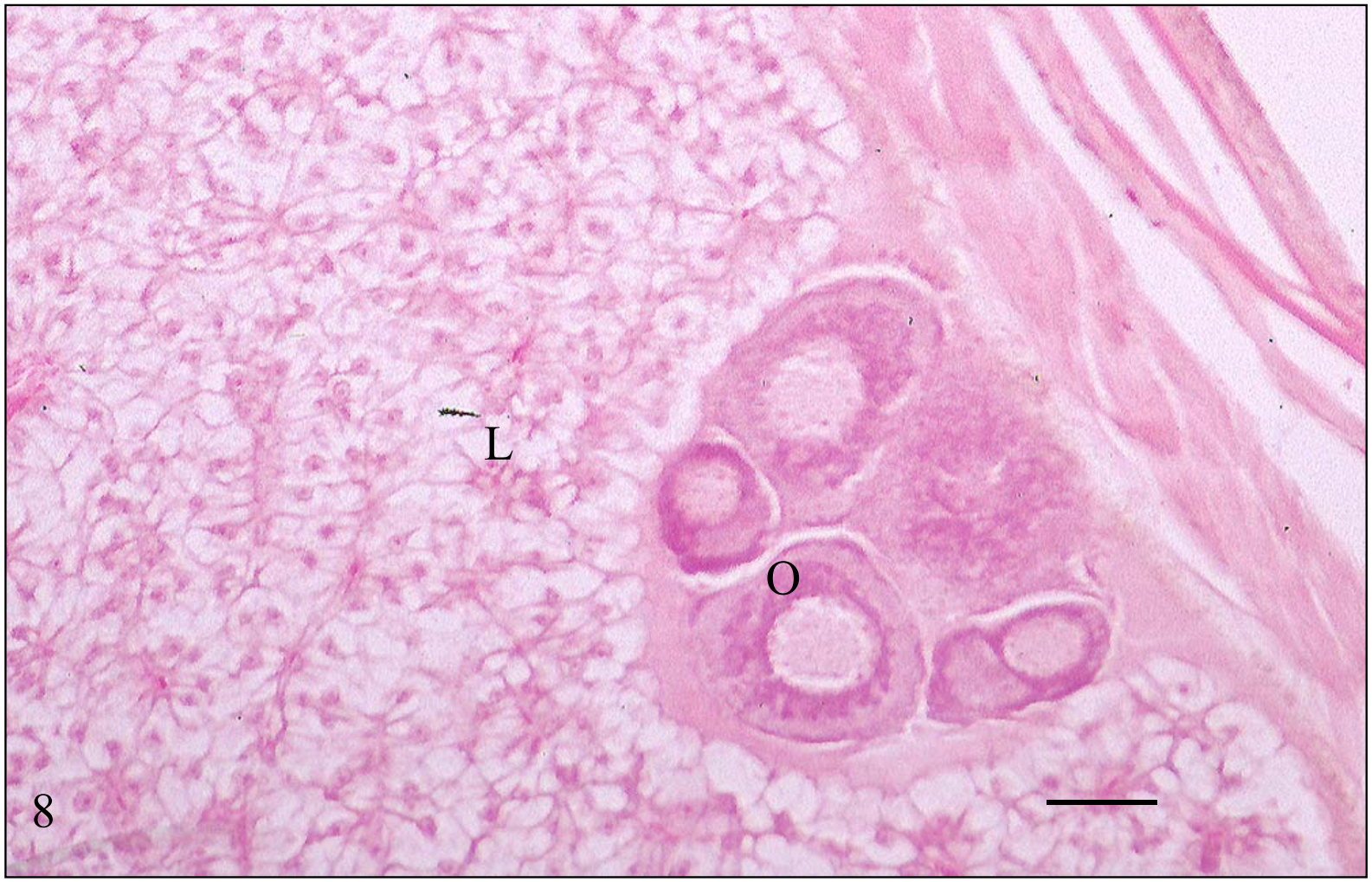


Figure 8. Ectopic ovary (O) in the liver (L) of medaka exposed to 100 ppm chloroform as embryos and evaluated after 6 months. H&E, bar=25 μm.

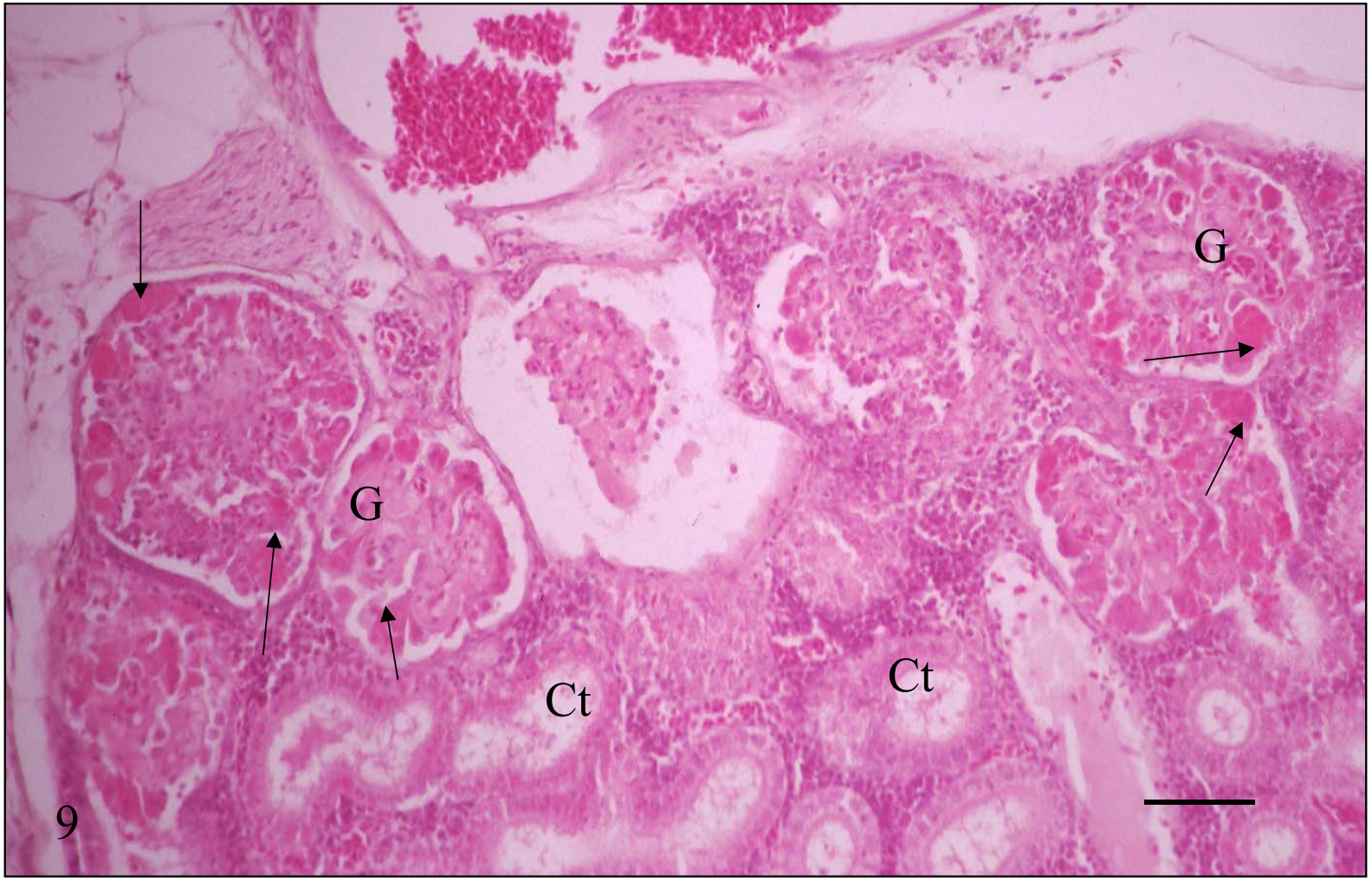


Figure 9. Glomerular (G) hyaline (arrows) mass and cystic tubules (Ct) in a kidney of a control medaka evaluated after 12 months. H&E, bar=30 μm.

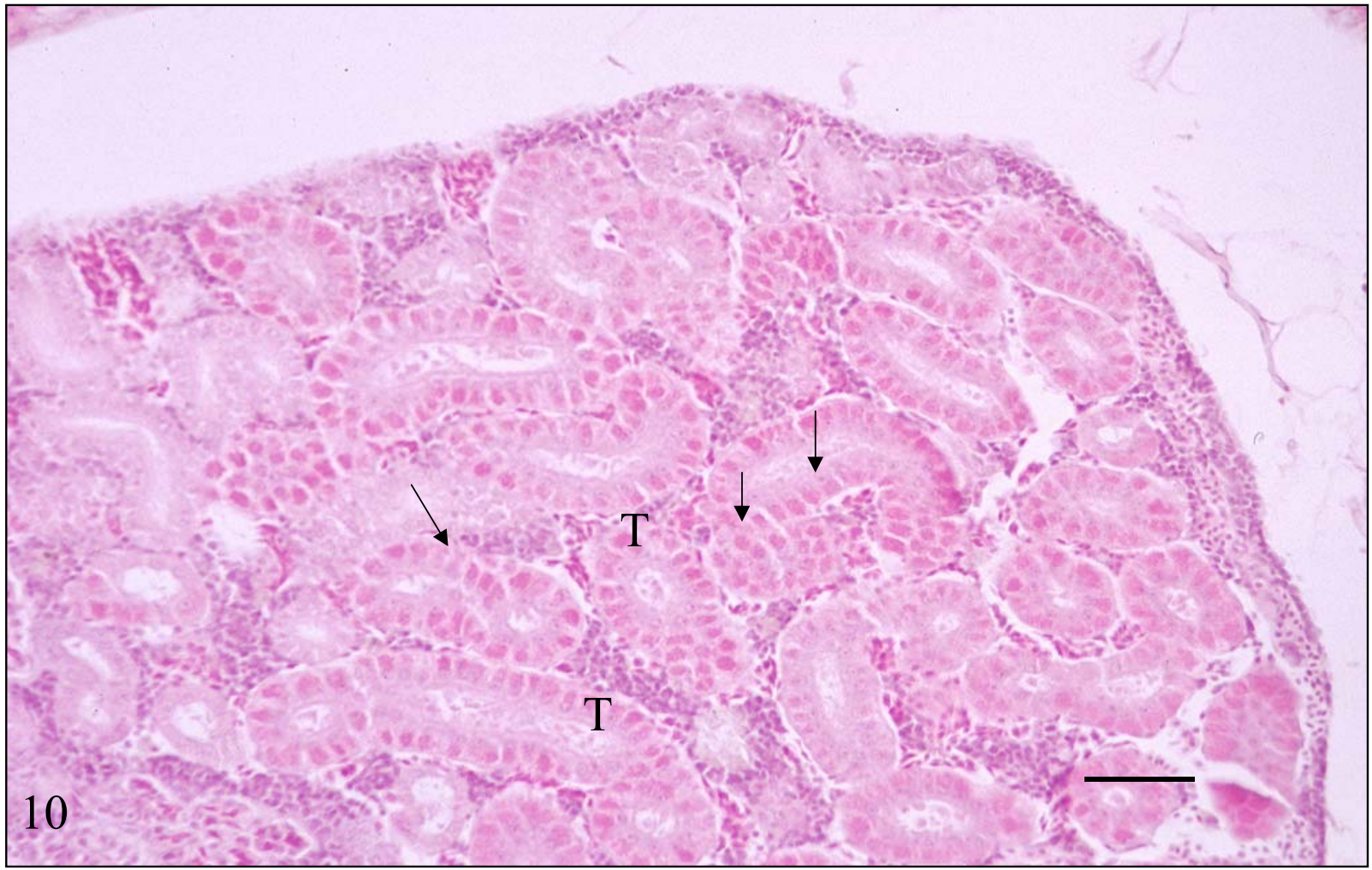


Figure 10. Hyaline globules (arrows) in the tubular epithelium (T) of the kidney of a control medaka evaluated after 12 months. H&E, bar=25 μm.

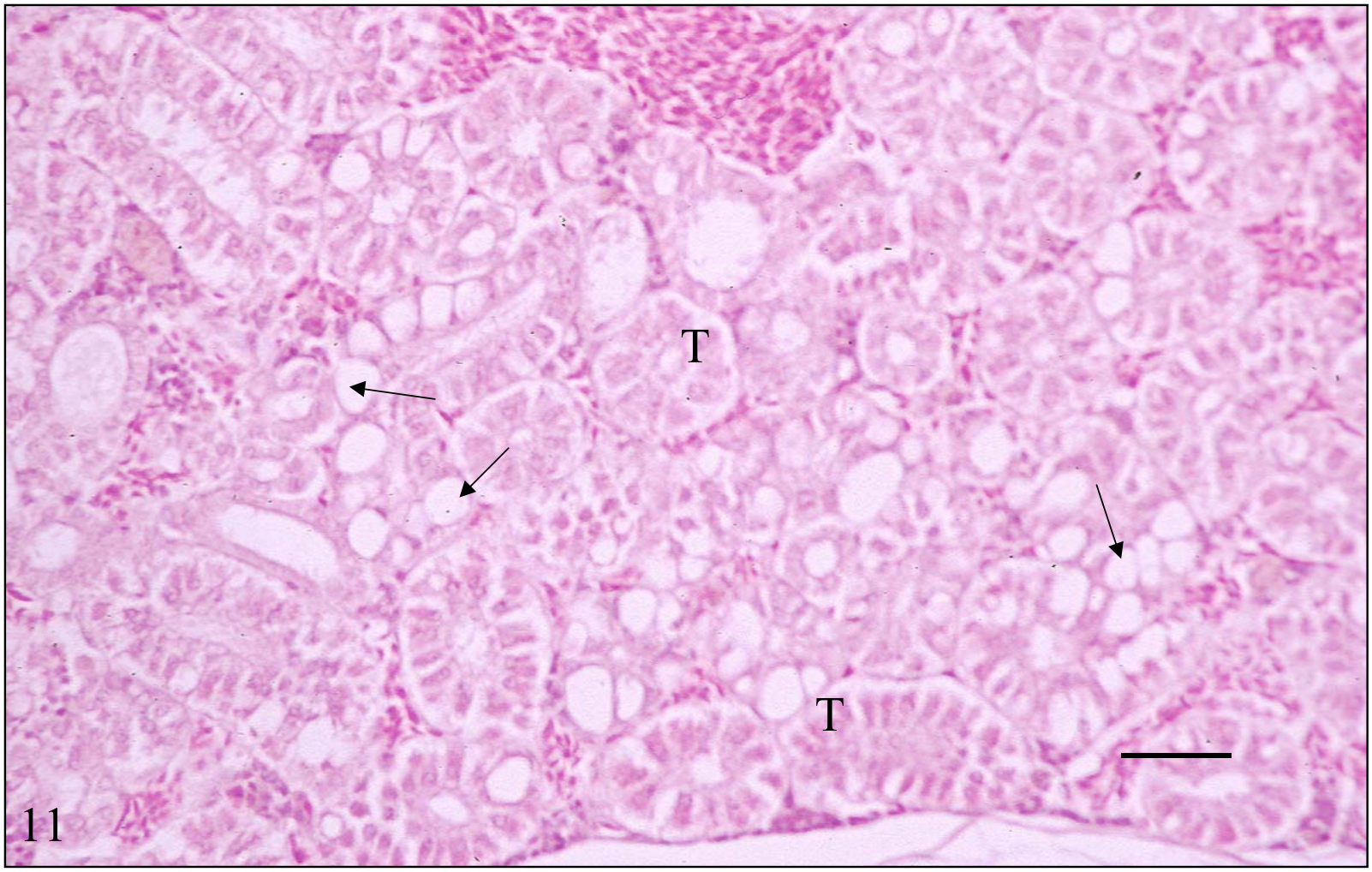


Figure 11. Vacuoles (arrows) in the tubular epithelium (T) of the kidney of a medaka exposed as embryos to 50 ppm chloroform for 10 days and evaluated after 12 months. H&E, bar=25 μm.

Glomerular lesions observed were accumulation of hyaline material in the glomerular tuft (Figure 9) and dilated Bowman's space (Figure 9), which often contained exudate. Interstitial lesions observed were granulomas and low incidence of interstitial tissue necrosis (Table 2). Granulomas were often associated with infectious agents. Because we observed protozoan parasites in association with granulomas, we speculate that the parasite is the cause for granulomas in our study. One case of lipoma was observed in the kidney of a medaka exposed to 25 ppm chloroform and evaluated after 12 months.

Gonads

Ovaries were classified into three categories (Table 3) based on the histomorphology of the oocytes during sexual maturation. They were 1) previtellogenic oocytes (Figure 12), characterized by perinucleolar and alveoli oocytes; 2) early vitellogenic, characterized by predominantly previtellogenic oocytes and few vitellogenic oocytes; and 3) mature ovary (Figure 13), containing various stages of oocytes, predominantly vitellogenic and attretic follicles. Cystic ovary (Figure 14) is a pathological condition characterized by several fluid follicles (ruptured or intact), surrounded by the epithelium of the oviduct. Cystic ovary was found in fish examined after 12 months of grow-out, including control fish.

Testes were grouped into two categories, (Table 3) immature or mature. Immature testes (Figure 15) were characterized by a uniform population of cells with no stages of spermatogenesis. Mature testes (Figure 16) were characterized by spermatogenesis including presence of sperms.

Ectopic gonad was defined as the occurrence of gonadal tissue outside its normal location, including in the anterior abdominal cavity.

In chloroform-exposed fish, delayed sexual maturation was observed at all concentrations, including controls. Immature testes were found only in fish exposed to chloroform and examined after 6 months. None of the 12-month-old fish had immature testes except one control fish. The incidence of immature testes in chloroform-exposed fish examined after 6 months was higher compared to the controls. In female fish, immature ovaries occurred in all groups including controls. However, the incidence was higher in chloroform-exposed fish than controls. In addition, one fish exposed to 100 ppm and examined after 6 months had previtellogenic oocytes in the liver and the anterior abdominal cavity. One control female fish had ectopic testicular tissue in the anterior abdominal cavity.

Thyroid

Thyroid tissue is normally found as scattered follicles (Figure 17) around the aorta and gill arch. In this study, approximately 5 to 10 small follicles in a section is considered normal. The same number of follicles with an increase in size was considered hypertrophy. An increase in the number of follicles was considered follicular hyperplasia. Hypertrophy may or may not

Table 2

Incidence of Renal Lesions in 6- and 12-Month-Old Fish Exposed to Chloroform as Embryos

Lesions	ERS		DMSO	25 ppm		50 ppm		100 ppm	
	6M	12M	6M	6M	12M	6M	12M	6M	12M
Total N	60	60	60	60	77	60	62	60	32
Number examined	38	47	31	60	66	58	51	55	25
NAD	20	34	13	4	29	19	22	18	6
Cystic tubules	16	10	15	48	31	19	21	20	9
Tubular hyaline	4	1	1	30	0	26	0	24	1
Vacuoles in tubular epithelium	3	0	0	1	0	0	1	0	0
Tubular necrosis	0	0	0	2	0	0	0	1	0
Thyroidization	0	0	0	0	0	0	0	0	0
Glomerular hyaline	5	0	0	0	0	0	0	0	0
Dilated Bowman's space	2	1	2	11	5	1	4	0	1
Granuloma	0	4	1	4	11	3	8	0	13
I/S necrosis	0	0	0	0	1	0	0	0	0
Lipoma	0	0	0	0	1	0	0	0	0

ERS = embryo-rearing solution
DMSO = dimethyl sulfoxide
M = months
NAD = nothing adverse detected

Table 3

Incidence of Gonadal Lesions in 6- and 12-Month-Old Fish Exposed to Chloroform

Lesions	ERS		DMSO	25 ppm		50 ppm		100 ppm	
	6M	12M	6M	6M	12M	6M	12M	6M	12M
Total N	60	60	60	60	77	60	62	60	32
Number examined	13	21	19	50	46	51	47	50	16
Males N	2	6	4	20	11	19	9	20	5
Females N	11	15	15	30	35	32	38	30	11
Immature testis	0	1	0	9	0	6	0	11	0
Mature testis	2	5	0	11	11	13	9	9	5
Granuloma	0	0	0	0	0	0	1	0	2
Previtellogenic	6	1	4	19	16	12	10	17	0
Early vitellogenic	5	1	1	8	7	2	4	4	0
Mature ovary	0	13	10	3	12	18	24	9	11
Cystic ovary	0	4	0	0	3	0	5	0	0
Granuloma	0	1	0	0	0	0	1	0	0
Ectopic gonad	1	0	0	0	0	0	0	1	0

ERS = embryo-rearing solution
DMSO = dimethyl sulfoxide
M = months

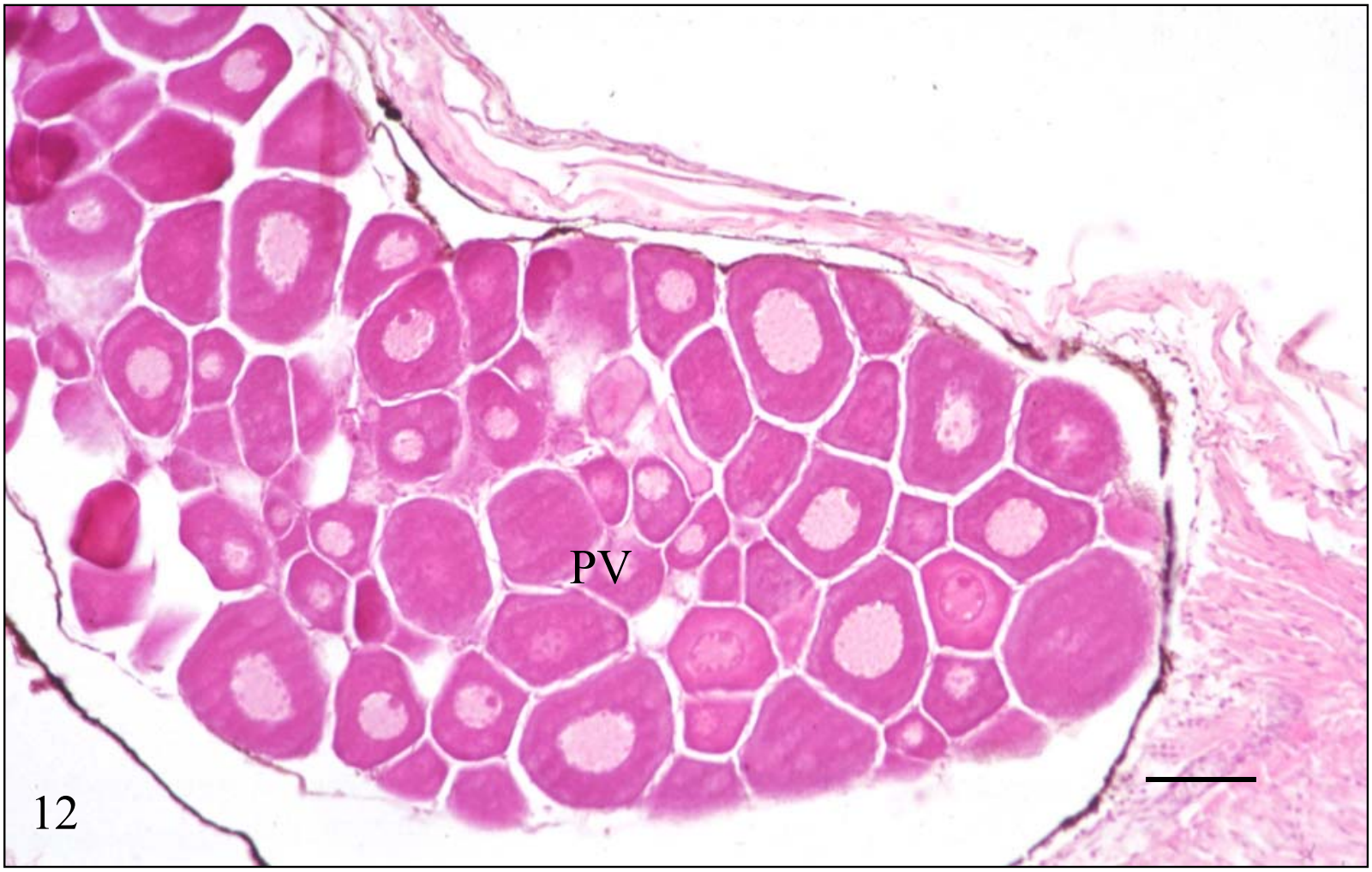


Figure 12. Previtellogenic oocytes (Pv) in a medaka exposed as embryos to 25 ppm chloroform for 10 days and evaluated after 6 months. H&E, bar=50 μm.

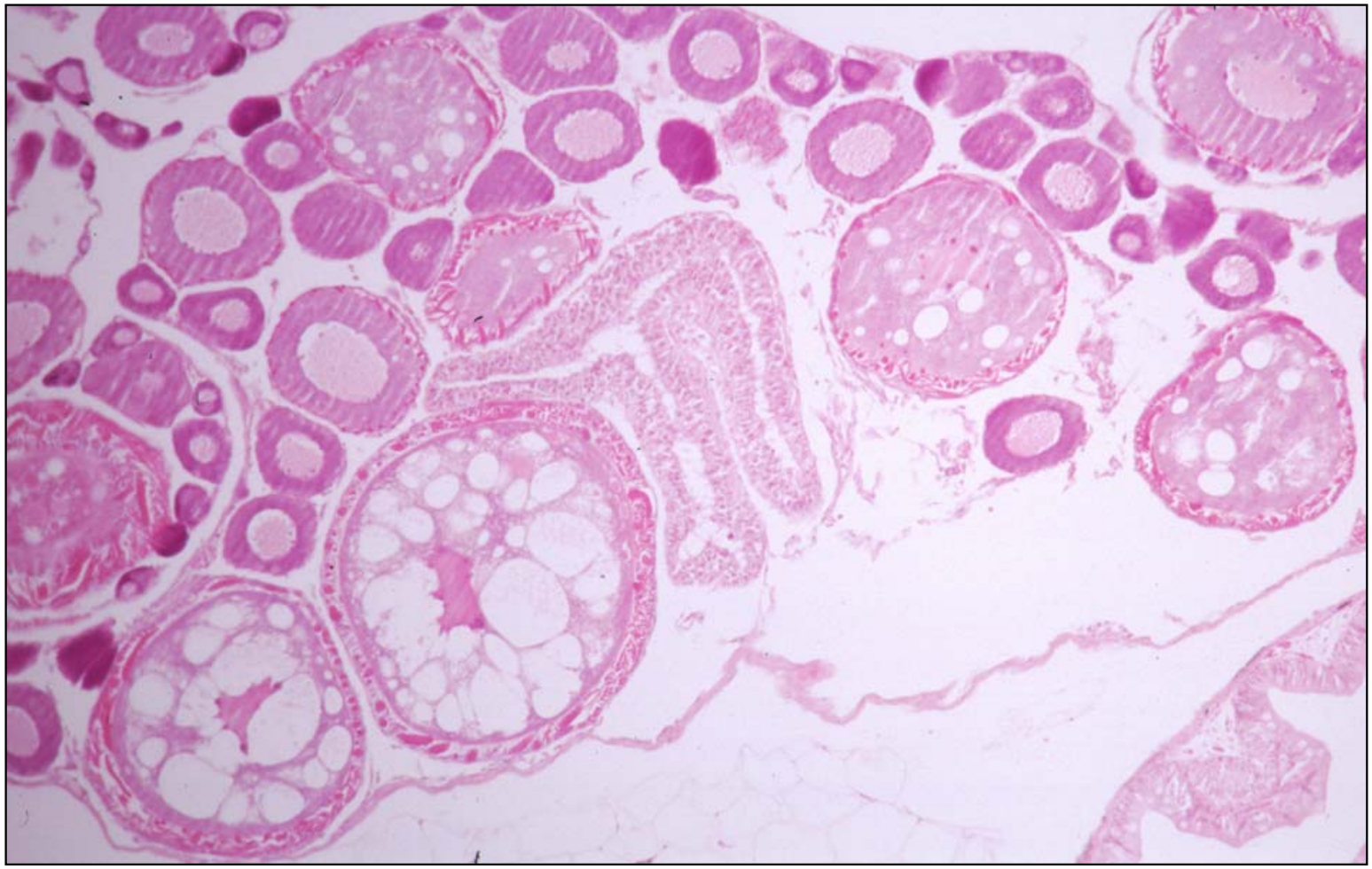


Figure 13. Mature ovary of a control medaka. Vitellogenic oocytes (V), previtellogenic oocytes (Pv), and atretic follicles (At). H&E, bar=50 μm.

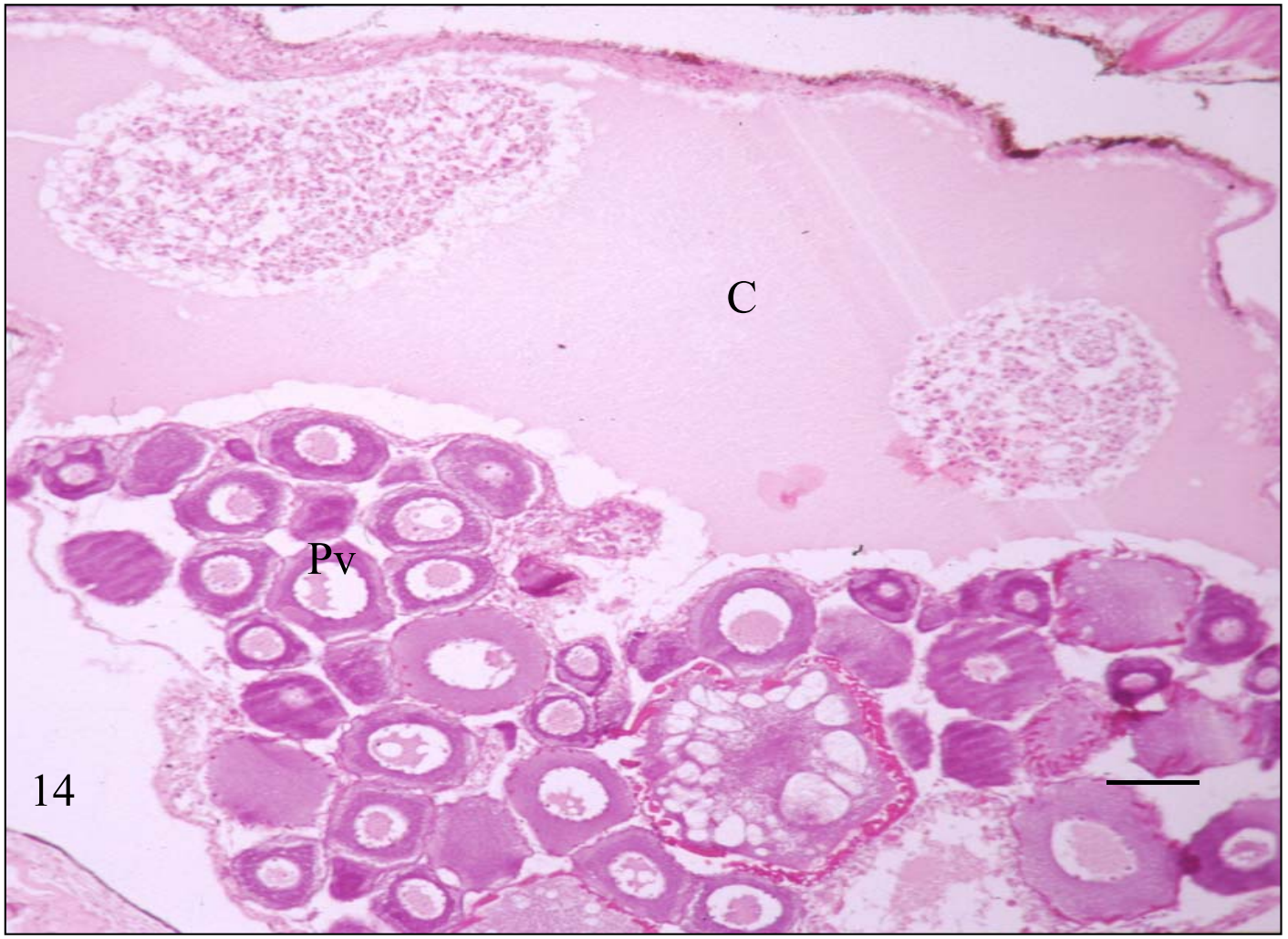


Figure 14. Cystic ovary (C) of a medaka exposed as embryos to 50 ppm chloroform for 10 days and evaluated after 12 months. Note the various stages of oocytes (F). H&E, bar=40 μm.

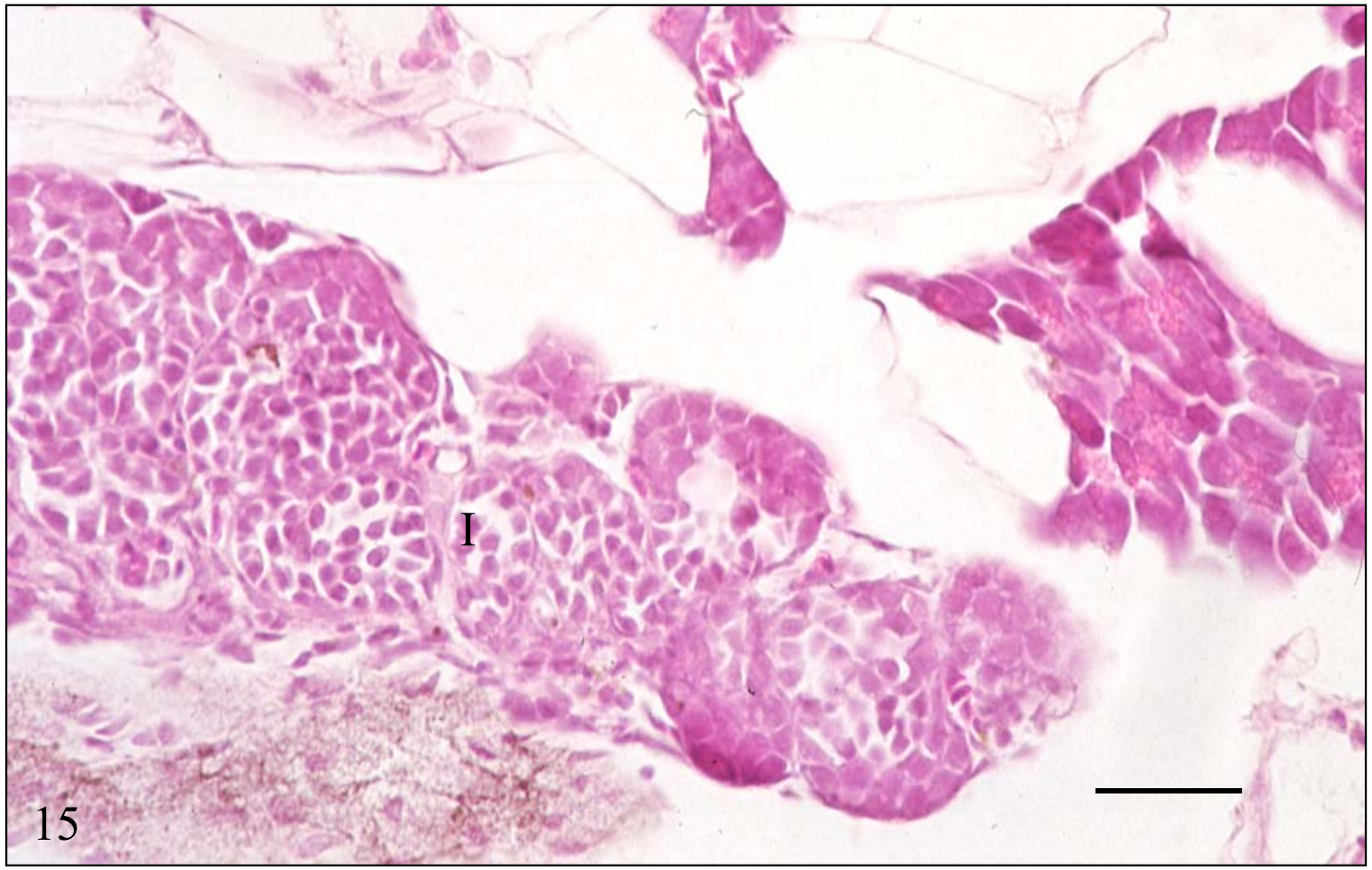


Figure 15. Immature testis (I) in a medaka exposed as embryos to 100 ppm chloroform for 10 days and evaluated after 6 months. H&E, bar=35 μm.

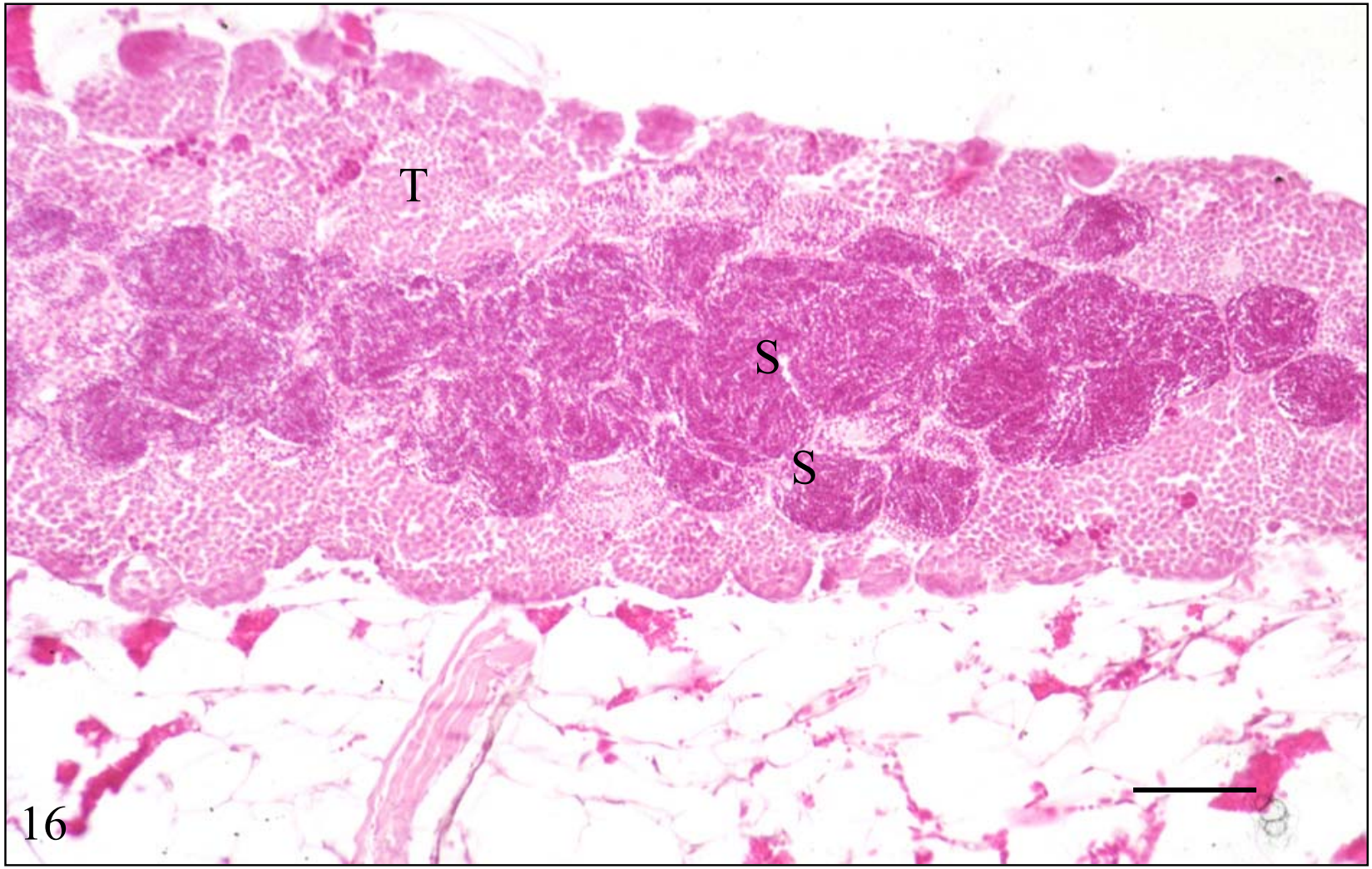


Figure 16. Mature testis (T) of a control medaka showing abundant sperms (S). H&E, bar=30 μm.

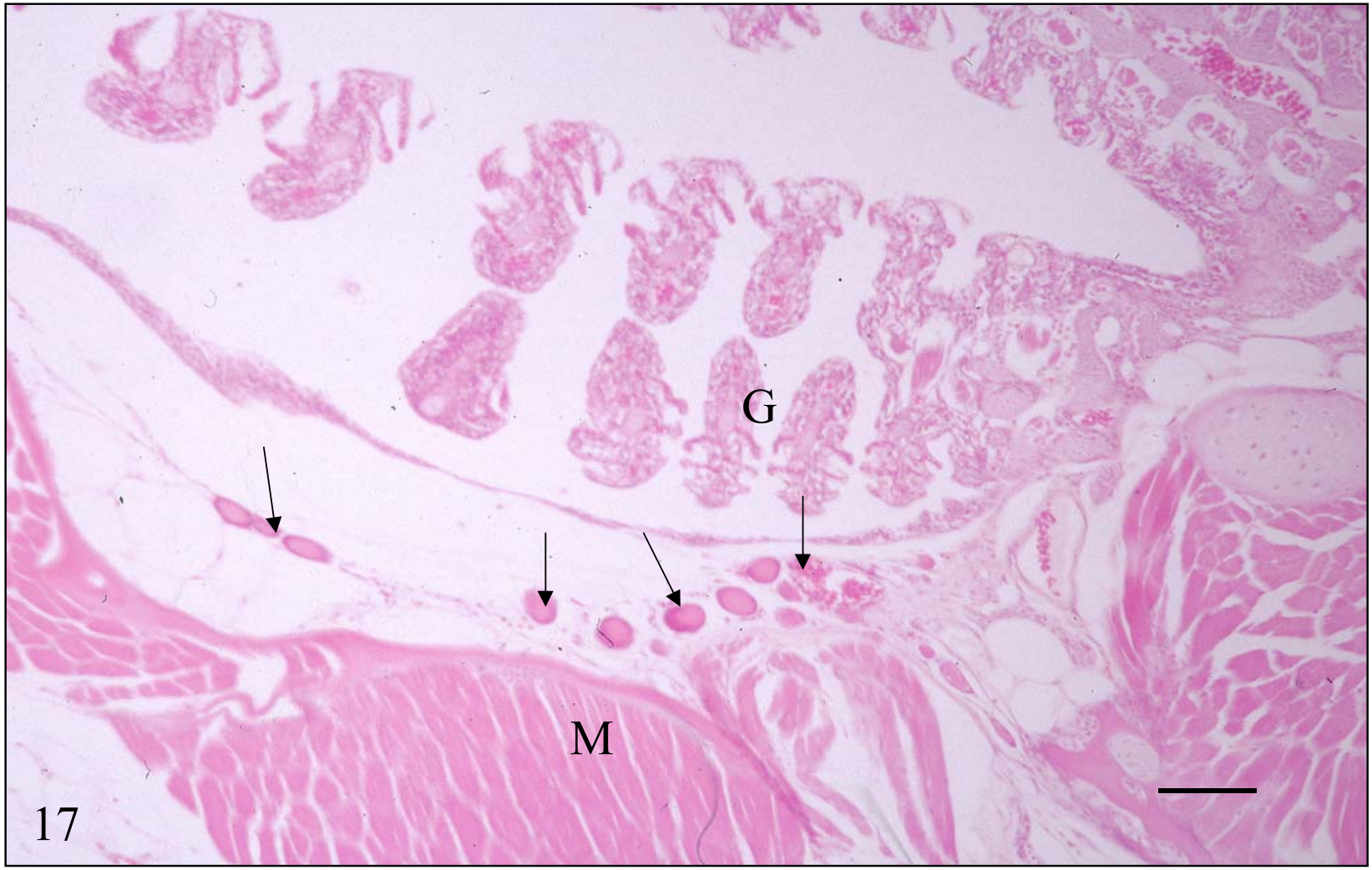


Figure 17. Thyroid follicles (arrows) in a control medaka. Gills (G) and Muscle (M). H&E bar=45 μm.

be associated with hyperplasia. There was an apparent increase in the follicular hyperplasia (Figure 18) in fish exposed to chloroform and evaluated after 12 months compared to those checked after 6 months.

Bromoform Studies

Lesions were grouped into various categories as described in the chloroform studies. Incidence of thyroid lesions (Table 4), liver lesions (Table 5), kidney lesions (Table 6), gonadal lesions (Table 7) and thyroid lesions (Table 8) are tabulated. Liver lesions in bromoform-exposed fish followed incidence patterns similar to those observed in chloroform-exposed fish.

Incidence of hepatic cysts were higher in bromoform-exposed fish compared to the control fish. One fish exposed to 50 ppm bromoform and examined after 6 months had cholangioma.

High incidence of cystic tubules occurred in all groups of fish including controls. Thyroidization of tubules occurred in two fish exposed to 50 ppm and examined after 6 months. However, chloroform-exposed fish did not have this lesion. A higher incidence of immature testis and ovary occurred in 6-month-old fish exposed to bromoform at 10 ppm.

Follicular hyperplasia was found in fish from all groups including controls. Thyroid follicular hyperplasia occurs in aged fish (Hoover, 1984), iodine deficiency (Gaylord and Marsh, 1912), poor water quality (Nigrelli and Ruggieri, 1974), accumulation of excreta and metabolites in the water (Nigrelli and Ruggieri, 1974), high temperature (Hoover, 1984), or changes in estrogen levels (Barrington and Matty, 1954). In our study, the water quality parameters (temperature, ammonia, nitrite) were within acceptable limits. Also, high incidence of follicular hyperplasia occurred in 12-month-old fish including controls compared to 6-month-old fish. We reserve making definitive conclusions until all the samples are evaluated and statistically analyzed.

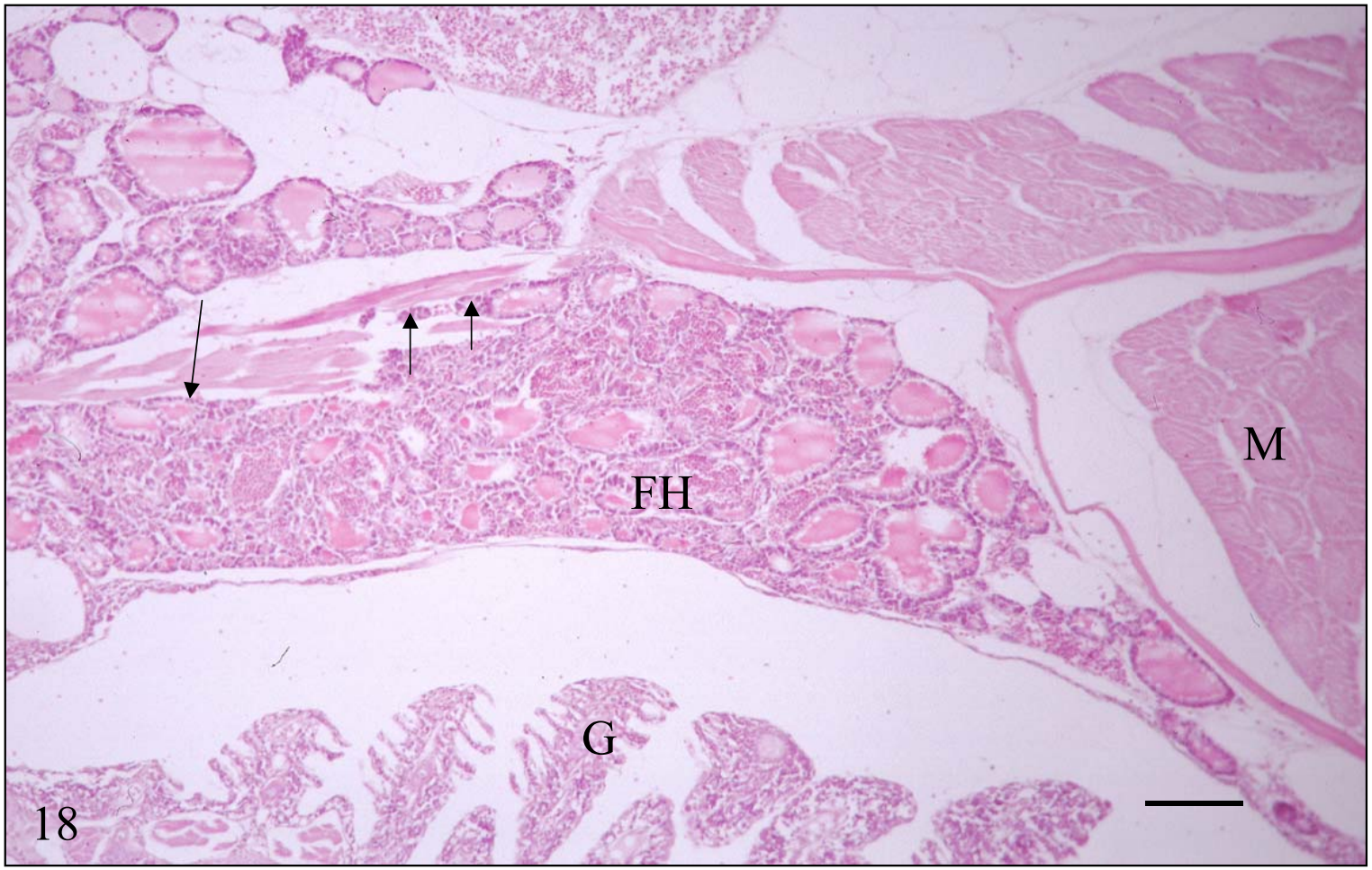


Figure 18. Follicular hyperplasia (Fh) of the thyroid in a medaka exposed as embryos to 50 ppm chloroform for 10 days and evaluated after 12 months. Proliferating follicles are infiltrating(arrows) adjoining muscle (M). Gills (G). H&E, bar=45 μm.

Table 4

Incidence of Thyroid Lesions in 6- and 12-Month-Old Fish Exposed to Chloroform

Lesions	ERS		DMSO	25 ppm		50 ppm		100 ppm	
	6M	12M	6M	6M	12M	6M	12M	6M	12M
Total N	60	60	60	60	77	60	62	60	32
Number examined	56	47	54	59	77	57	61	58	29
NAD	51	23	37	55	21	45	19	53	13
Hypertrophy only	1	7	7	2	19	3	14	1	8
Follicular hyperplasia	4	16	9	2	37	9	38	4	8
Infiltration into opercular muscles	1	5	1	0	8	0	10	0	2

ERS = embryo-rearing solution
DMSO = dimethyl sulfoxide
M = months
NAD = nothing adverse detected

Table 5

Incidence of Liver Lesions in 6- and 12-Month-Old Fish Exposed to Bromoform

Lesions	ERS		DMSO	10 ppm		50 ppm
	6M	12M	6M	6M	12M	6M
Total N	60	60	60	61	60	20
Number examined	52	42	47	61	26	20
NAD	45	26	39	15	9	2
Degenerative Lesions						
Excessive accumulation	2	0	1	31	5	14
Fatty change	1	3	2	1	1	2
Hepatic cysts	5	4	5	19	7	10
Spongiosis hepatis	0	4	1	7	3	4
Biliary cyst	0	0	0	1	0	0
Granuloma	0	0	0	0	10	2
Necrosis	0	4	0	0	0	8
Non-neoplastic Proliferative Lesions						
Bile duct hyperplasia	0	0	0	0	0	1
Altered foci	0	0	0	3	0	2
Clear cell foci	0	1	0	1	1	2
Vacuolated cell foci	0	2	0	0	0	2
Neoplasms						
Cholangioma	0	0	0	0	0	1
Miscellaneous						
Ectopic gonad	0	0	0	0	0	0

ERS = embryo-rearing solution
DMSO = dimethyl sulfoxide
M = months
NAD = nothing adverse detected

Table 6

Incidence of Kidney Lesions in 6- and 12-Month-Old Fish Exposed to Bromoform

Lesions	ERS		DMSO	10 ppm		50 ppm
	6M	12M	6M	6M	12M	6M
Total N	60	60	60	61	60	20
Number examined	38	47	31	55	17	20
NAD	20	34	13	18	7	3
Cystic tubules	16	10	15	25	7	15
Tubular hyaline	4	1	1	13	1	2
Vacuoles in tubular epithelium	3	0	0	0	0	1
Tubular necrosis	0	0	0	0	0	1
Glomerular hyaline	5	0	0	2	0	0
Dilated Bowman's space	2	1	2	3	0	3
Thyroidization	0	0	0	0	0	2
Granuloma	0	4	1	3	3	2
I/S necrosis	0	0	0	0	0	0
Lipoma	0	0	0	0	0	0

ERS = embryo-rearing solution
DMSO = dimethyl sulfoxide
M = months
NAD = nothing adverse detected

Table 7

Incidence of Gonadal Lesions in 6- and 12-Month-Old Fish Exposed to Bromoform

Lesions	ERS		DMSO	10 ppm		50 ppm
	6M	12M	6M	6M	12M	6M
Total N	60	60	60	61	60	20
Number examined	13	21	19	45	20	15
Males N	2	6	4	10	3	7
Females N	11	15	15	35	17	8
Immature testis	0	1	0	7	0	0
Mature testis	2	5	0	3	3	7
Granuloma	0	0	0	0	0	0
Previtellogenic	6	1	4	16	0	1
Early vitellogenic	5	1	1	1	0	0
Mature ovary	0	13	10	18	17	7
Cystic ovary	0	4	0	0	8	0
Granuloma	0	1	0	0	0	0
Ectopic gonad	1	0	0	0	0	0

Table 8

Incidence of Thyroid Lesions in 6- and 12-Month-Old Fish Exposed to Bromoform

Lesions	ERS		DMSO	10 ppm		50 ppm
	6M	12M	6M	6M	12M	6M
Total N	60	60	60	61	60	20
Number examined	56	47	54	55	27	20
NAD	51	23	37	47	8	12
Hypertrophy only	1	7	7	7	0	3
Follicular hyperplasia	4	16	9	1	19	5
Infiltration into opercular muscles	1	5	1	0	4	0

ERS = embryo-rearing solution
DMSO = dimethyl sulfoxide
M = months
NAD = nothing adverse detected

References

Barrington, EJ., and Matty, Aj. 1954. Seasonal variation in the thyroid gland of the minnow, Phoxinus phoxinus L., with some observations on the effect of temperature. Proc Zool Soc (London) 124: 89-95.

Boorman, GA., Botts, S., Bunton, TE., Fournie, JW., Harshbarger, JC., Hawkins, WE., Hinton, DE., Jokinen, MP., Okihiro, MS., and Wolfe, MJ. 1998. Diagnostic criteria for degenrative, inflammatory, proliferative nonneoplastic, and neoplastic liver lesions in medaka (Oryzias latipes): consensus of National Toxicology Program Pathology Working Group. Toxicologic Pathology 25: 202-210.

Gaylord, HR., and Marsh, MC. 1912. Carcinoma of the thyroid in the salmonid fishes. Bull Bureau Fish 32: 363-524.

Hoover, K. 1984. Hyperplastic thyroid lesions in fish. NCI Monograph 65: 275-289.

Kirchen, R.F., and West, W. 1979. The Japanese medaka its care and development, Burlingame, NC: Carolina Biological Supply Company, p. 36.

Nigrelli, RF., Ruggieri, GD. 1974. Hyperplasia and neoplasia of the thyroid in marine fishes. Mt. Sinai J Med (NY) 41: 283-293.

Rugh, R. 1962. Experimental embryology techniques and procedures, third ed., Minneapolis, MN: Burgess Publishing Co.

Wolf, MJ. 1994. Pathology assessment issues. In: Compendium of the FY 1990 & FY 1992. Research reviews for the research methods branch, U.S. Army Medical Research & Development Laboratory, Fort Detrick, Frederick, MD. HS. Gardner, Jr., Compiler, Technical Report 9404,
p. 83-87.

PROGRESS REPORT: THE HEPATOTOXIC INTERACTION OF CHLOROFORM, BROMODICHLOROMETHANE, BROMOFORM, AND CHLORODIBROMOMETHANE

J.E. Simmons[1], C. Gennings[2], A. McDonald[1], W.R. Hartley[3], A. Thiyagarajah[3], L.K. Teuschler[4], Y.M. Sey[1]

[1]NHEERL, U.S. EPA, RTP, NC
[2]Virginia Commonwealth University, Richmond, VA,
[3]Tulane University, Environmental Diseases Prevention Research Center, New Orleans, LA,
[4]NCEA, U.S. EPA, Cincinnati, OH

September 1999

Abstract

Chlorination of drinking water has led to remarkable improvements in public health, dramatically reducing outbreaks of water-borne diseases. Chlorination of water containing organic matter results in the production of a variety of halogenated by-products, including the trihalomethanes (THMs), chloroform ($CHCl_3$), bromodichloromethane (BDCM), chlorodibromomethane (CDBM), and bromoform ($CHBr_3$). Given the near ubiquitous human exposure to the four THMs and the concern over their toxicity as single chemicals, assessment of the nature (additive, synergistic, or antagonistic) of the interaction between them will provide useful information in assessing the human health risks associated with disinfection of water. A threshold additivity model was developed for this purpose (Gennings et al., 1997); in this method, the response at specific mixture concentrations of interest is predicted under dose addition by data from the single chemical dose-response curves for each chemical in the mixture. The concentrations of the four THMs mixtures tested at total doses of 0.436 and 0.872 mmol/kg/day were based on the average seasonal proportions at 35 water treatment facilities (Krasner et al., 1989). Female CD-1 mice, ~65 to 70 days old, were gavaged daily for 14 days with this mixture in 10% Alkamuls®. Based on the single chemical dose-response curves generated in our lab, both the serum sorbitol dehydrogenase (SDH) and alanine aminotransferase (ALT) levels predicted by the threshold additivity model for both mixtures fell within 95% prediction intervals of what the response should be if dose-addition is a good assumption. The closeness of the predicted mixture responses and the observed mixture response indicates that the four chemicals are dose-additive at the mixture combination tested.

Introduction

Chlorination of drinking water has led to remarkable improvements in public health, dramatically reducing outbreaks of water-borne diseases. However, the disinfection of water containing organic matter results in the production of a variety of halogenated by-products, including the trihalomethanes (THMs), chloroform ($CHC1_3$), bromodichloromethane (BDCM), chlorodibromomethane (CDBM), and bromoform ($CHBr_3$). At high laboratory concentrations, these THMs have been shown individually to be either carcinogenic or to cause target-organ toxicity. However, there is relatively little expectation that exposure to individual THMs at levels found in drinking water is likely to result in significantly increased human health risk. Thus, it is important to assess the nature (additive, synergistic, or antagonistic) of any toxicologic interactions among them to provide useful information in assessing the human health risks associated with disinfection of water.

To address this issue, the objectives of this research were to:

- provide data for the development of the threshold additivity model
- develop an appropriate experimental design for the collection of data on mixtures to be analyzed by the threshold additivity model
- develop an understanding of the toxicity (i.e., the potency and nature of the toxic interaction) of the four THMs formed during disinfection of drinking water

Specifically, this project examined the hepatotoxicity in mice of each of the four THMs, singly and together in mixtures. The experimental animal was selected to determine whether single chemical data from the published literature could be used to construct the expected response of a chemical mixture. Female CD-1 mice were selected because only one report (Munson et al., 1982) was found in the published literature where the hepatotoxicity of each of the four THMs was investigated in the same laboratory following subacute administration in an aqueous vehicle. The liver was identified as one of the critical target organs based on information found in U.S. EPA's Integrated Risk Information System (1999). Serum enzymes sorbitol dehydrogenase (SDH), alanine aminotransferase (ALT), and aspartate aminotransferase (AST) were measured as markers for hepatotoxicity and statistically analyzed for interaction effects. Histopathology was also performed for comparison with the serum enzyme data.

Assessment of the interaction of chemicals is not an easy task. A significant concern in mixtures research regards the quality of the experimental designs and statistical analyses used to determine the validity of the assumption of additivity. For non-cancer health effects, the method of dose addition has been recommended by the U.S. EPA (1986). Dose addition for two chemicals may be addressed by Berenbaum's equation (1985, 1989), where, under dose-additivity, the contours of constant response of the dose-response surface are planar (i.e., linear for two chemicals). These planar contours are given by:

$$\sum_{i=1}^{c} \frac{x_i}{X_i} = 1. \quad (1)$$

where X_1, X_2 ...X_c are isoeffective or equally toxic doses of the c individual chemicals such that each produces the same magnitude of effect as the mixture, and the x_1, x_2 ...x_c are the dose levels in the mixture that would produce the same effect. Although the traditional isobologram is a convenient way to interpret the results of the simplest form of dose addition, i.e., mixtures involving two chemicals, determining whether chemicals are dose additive becomes increasingly difficult as the number of chemicals increases. To address this need, as well as the need for efficient experimental designs that require fewer treatment groups and animals than traditional full-factorial designs, a threshold model for mixtures under the assumption of dose-additivity has been developed (Gennings and Carter, 1995; Gennings et al., 1997). In this method, data from the single-chemical dose-response curves for each chemical in the mixture are used to estimate the response, under dose addition, at specific mixture combinations of interest. These specific mixtures may then be tested in the laboratory and the empirical results compared with the model predictions to assess the joint action of the chemicals (i.e., additivity, synergism, or antagonism).

Experimental Methods

To estimate the response under dose addition, single-chemical dose-response curves are needed. To this end, each THM was assessed in a separate experiment. Female CD-1 mice (Charles River Breeding Laboratory, Raleigh, NC) were gavaged daily, for 14 days with one of the four THMs, in 10% Alkamuls® EL-620. Dosages were 0, 0.152, 0.305, 0.76, 1.52, and 3.05 mmol/kg/d for $CHCl_3$, $CHBr_3$, and BDCM. Because of an error, the dosages for CDBM were 0, 0.304, 0.610, 1.52, 3.04, and 6.10 mmol/kg/d. Groups of 8 to 20 animals were gavaged in the morning with dosing solutions that were made fresh daily. The gavage volume was 10 ml/kg. On the morning following the 14th day of dosing, anoxia was induced by carbon dioxide. Mice were bled from the abdominal aorta. Serum was prepared and stored at -80°C until analyzed for the activity of SDH, ALT, and AST by automated procedures (COBAS Fara II).

Tissues were preserved in 10% phosphate buffered formalin for shipping. Serial sections, 5 μm thick, were cut from paraffin-embedded tissue. Two of six slides of the liver were stained with hematoxylin and eosin and examined blind (without knowledge of dose group) by brightfield microscopy. Centrilobular necrosis was characterized as none (no necrotic hepatocytes seen), mild (less than 10% of the hepatocytes involved), mild to moderate (less than 50% of the hepatocytes involved), moderate (approximately 50% of the hepatocytes involved), moderate to severe (more than 50% of the hepatocytes involved), and severe (more than 75% of the hepatocytes involved).

The same experimental methods were used for the 4-THM mixtures. The ratios of the chemicals in the mixture were based on the average seasonal proportions at 35 water treatment facilities (Krasner et al., 1989). The proportions of each component were 0.65 $CHCl_3$, 0.01 $CHBr_3$, 0.24 BDCM, and 0.10 CDBM, at two total dose combinations. The dose combinations

were: 0.568, 0.012, 0.208, and 0.084, mmol/kg/d for $CHCl_3$, $CHBr_3$, BDCM, and CDBM, respectively, for a total dosage of 0.872 mmol/kg/d, and 0.284, 0.006, 0.104, and 0.042 mmol/kg/d, respectively, for a total dosage of 0.436 mmol/kg/d. Concurrently, single-chemical data points were included for comparison with the single-chemical dose-response curves: $CHCl_3$ at 2.10 mmol/kg/d (250 mg/kg/d), $CHBr_3$ at 1.00 mmol/kg/d (250 mg/kg/d), BDCM at 1.52 mmol/kg/d (250 mg/kg/d), and CDBM at 1.20 mmol/kg/d (250 mg/kg/d).

Data Analysis Methods

The single chemical data were analyzed by analysis of variance (ANOVA). When data violated the assumptions of normality and equality of variances, the data were transformed prior to analysis. When the overall ANOVA was significant, Dunnett's test was used to compare the treatment groups to the control. The criterion of significance was $p \leq 0.05$.

For the histopathology data, Chi-square tests were conducted for linear trends in dose-response. The Fisher Exact test was applied to the hepatic centrilobular incidence data for determination of the no-observed-adverse-effect level (NOAEL), the highest dose at which statistical significance was not achieved.

The mixture data were analyzed according to the methodology developed and described in Gennings et al. (1997) (see Section 5 of this report). This threshold additivity model was used to predict the response under dose addition at the particular mixtures tested. With increasing dose-response curves, if the observed response is greater than the predicted response, the response is considered synergistic (effect greater than expected under additivity), whereas if the observed response is less than the predicted response, the response is considered antagonistic (effect less than expected under additivity).

Results

Single Chemicals

Hepatotoxicity was observed in all four THM studies. Table 1 gives the lowest statistically significant doses for the serum endpoint and liver weight data. As shown in Table 1, statistically significant differences in SDH generally occurred at doses less than or equal to those for ALT and AST, with ALT as the next most sensitive indicator.

Liver histopathology in the controls was generally without abnormal deviation. However, there was a low prevalence of mild vacuolation with no zonal pattern in two of the experiments. All THMs showed a dose-related increase in incidence and severity of hepatocellular necrosis. The general patterns of vacuolation and necrosis were similar for the THMs, with CDBM producing the most severe lesions. Vacuolation was observed in Zone 2 (midzonal area) and Zone 3 (centrilobular area). The oxygen gradient and differential distribution of enzymes across zones may explain the zonal distribution of vacuolation and necrosis.

Table 1

Lowest Statistically Significant* Dosages

	$CHCl_3$	$CHBr_3$	BDCM	CDBM
	mg/kg/day (mmol/kg/d)			
Body weight (g)	NS	NS	500 (3.05)	NS
Relative liver weight ([liver/body] x 100)	181 (1.52)	77 (0.305)	125 (0.76)	63 (0.304)
SDH (IU/L)	91 (0.76)	192 (0.76)	50 (0.305)	127 (0.610)
AST (IU/L)	181 (1.52)	384 (1.52)	249 (1.52)	317 (1.52)
ALT (IU/L)	91 (0.76)	192 (0.76)	125 (0.76)	317 (1.52)

* Dunnett's test at $p \leq 0.05$; statistically significant differences from concurrent controls.
NS = not statistically significant

Chloroform: Relative liver weight and AST were increased at ≥1.52 mmol/kg/d. Significant increases in serum SDH and ALT occurred at ≥0.76 mmol/kg/d. Figure 1 illustrates the increases in SDH levels with increases in dose. Table 2 shows there were dose-related increases in incidence and severity of centrilobular necrosis. The incidence of centrilobular necrosis was: 1/12, 0/13, 8/15, 14/15, 13/13, and 12/12 for 0, 0.15, 0.305, 0.76, 1.52, and 3.05 mmol/kg, respectively. Mild to moderate necrosis occurred at 0.305 mmol/kg/d, which progressed in severity with dose to moderate to severe necrosis with severe Zone 2 vacuolation at the highest dose. The NOAEL for centrilobular necrosis is 0.152 mmol/kg/d.

BDCM: There were increases in SDH beginning at 0.305, in ALT and relative liver weight at ≥0.76, and in AST at ≥1.52 mmol/kg/d. Figure 2 illustrates the increases in SDH levels with increases in dose. Table 3 show the dose-related increase in incidence and severity of hepatic centrilobular necrosis. The incidence of centrilobular necrosis was: 0/10, 0/10, 1/11, 9/11, 10/10, and 8/8 for 0, 0.152, 0.305, 0.76, 1.52, and 3.05 mmol/kg/d, respectively. Liver pathological examination showed mild to moderate fatty changes at 0.152 and 0.305 mmol/kg/d, and mild centrilobular necrosis at 0.305 mmol/kg/d. At 0.76 mmol/kg/d, mild fatty changes and moderate centrilobular degeneration/necrosis occurred. At 1.52 mmol/kg/d, severe fatty changes and moderate to severe centrilobular necrosis were found. At 3.05 mmol/kg/d, there was severe centrilobular necrosis frequently accompanied by severe Zone 2 vacuolation. The NOAEL for centrilobular necrosis is 0.152 mmol/kg/d.

Bromoform: Relative liver weight increased at ≥0.305 mmol/kg/d. There were significant increases in serum SDH and ALT at ≥0.76 mmol/kg/d, and AST at ≥1.52 mmol/kg/d. Figure 3 illustrates the increases in SDH levels with increases in dose. Table 4 shows the dose-related increases in incidence and severity of centrilobular necrosis. The incidence for centrilobular necrosis was: 2/16, 0/14, 8/15, 14/16, 10/10, and 14/14 for 0, 0.152, 0.305, 0.76, 1.52, and 3.05 mmol/kg/d, respectively. At 0.305 mmol/kg/d, centrilobular necrosis was predominately mild to moderate, whereas at the two highest doses, moderate to severe cases occurred accompanied by Zone 2 vacuolation. The NOAEL for centrilobular necrosis is 0.152 mmol/kg/d.

CDBM: All mice died at 6.10 mmol/kg/d. Relative liver weight was increased at ≥0.304 mmol/kg/d. Serum enzymes (SDH at ≥0.610 mmol/kg/d, AST and ALT at ≥1.52 mmol/kg/d) were increased. Figure 4 illustrates the increases in SDH levels with increases in dose. Table 5 shows the dose-related increase in incidence and severity in fatty liver and centrilobular necrosis. The incidence of centrilobular necrosis was 0/13, 3/16, 17/18, 14/14, and 5/5 for 0, 0.304, 0.610, 1.52, and 3.04 mmol/kg/d, respectively. Severity of lesions were mild at the lowest dose, mild to moderate at 0.610 mmol/kg/d, and moderate and severe at the two highest doses. The NOAEL for centrilobular necrosis is 0.304 mmol/kg/d.

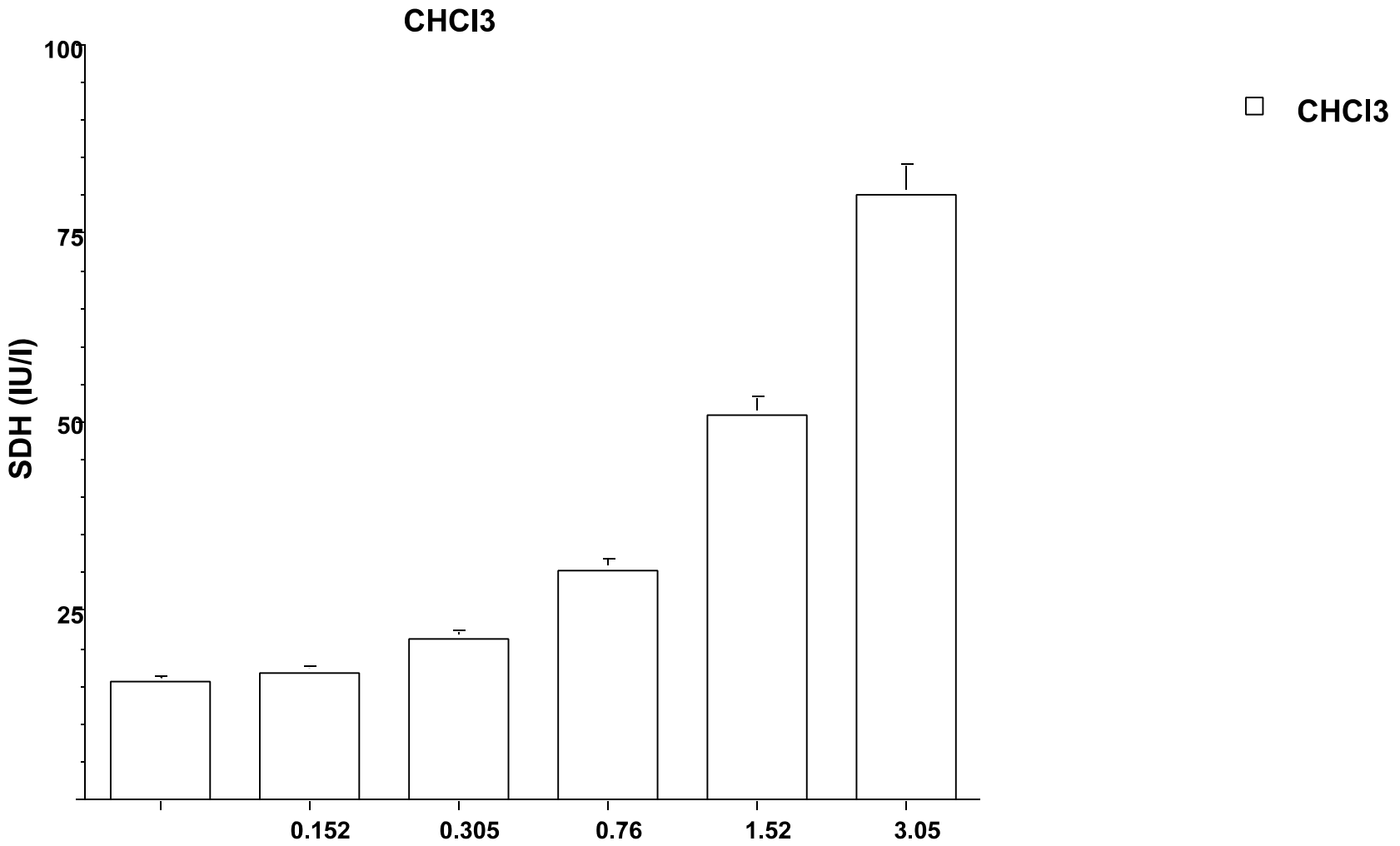


Figure 1
Dose-Related Changes in Mean SDH Levels for $CHCl_3$

Table 2. Incidence and Severity of Centrilobular Necrosis in Chloroform-Treated Mice

	Dose [mmol/kg/d (mg/kg/day)]					
	0	0.152 (18)	0.305 (36)	0.76 (91)	1.52 (181)	3.05 (364)
Severity	Percentage of animals					
No cenlob	92%	100%	47%	7%	0	0
mild	8%	0	26%	7%	8%	0
mild-mod	0	0	20%	7%	0	0
moderate	0	0	7%	53%	54%	33%
mod-severe	0	0	0	13%	23%	17%
severe	0	0	0	13%	15%	50%
Total[a]	8%[b]	0[c]	53%	93%	100%	100%
N	12	13	15	15	13	12

a — There was a significant dose-response ($p<0.0001$). b — There was one case of centrilobular necrosis in the control group. c — NOAEL = 0.152 mmol/kg/d (18 mg/kg/day) [Fisher Exact Test].

Table 3. Incidence and Severity of Centrilobular Necrosis in BDCM-Treated Mice

	Dose [mmol/kg/d (mg/kg/day)]					
	0	0.152 (25)	0.305 (50)	0.76 (125)	1.52 (249)	3.05 (500)
Severity	Percentage of animals					
No cenlob	100%	100%	91%	18%	0	0
mild	0	0	0	18%	0	0
mild-mod	0	0	0	9%	0	0
moderate	0	0	9%	55%	60%	0
mod-severe	0	0	0	0	30%	25%
severe	0	0	0	0	10%	75%
Total[a]	0	0	9%[b]	82%	100%	100%
N	10	10	11	11	10	8

a — There was a significant dose-response ($p<0.0001$).
b — NOAEL = 0.305 mmol/kg/d (50 mg/kg/day) [Fisher Exact Test].

Table 4. Incidence and Severity of Centrilobular Necrosis in Bromoform-Treated Mice

	Dose [mmol/kg/d (mg/kg/day)]					
	0	0.152 (38)	0.305 (77)	0.76 (192)	1.52 (384)	3.05 (771)
Severity	Percentage of animals					
No Cenlob	88%	100%	47%	12%	0	0
mild	12%	0	33%	19%	0	0
mild-mod	0	0	7%	0	0	0
moderate	0	0	13%	50%	50%	43%
mod-severe	0	0	0	19%	40%	7%
severe	0	0	0	0	10%	50%
Total[a]	12%[b]	0[c]	53%	88%	100%	100%
N	16	14	15	16	10	14

a — There was a significant dose-response ($p<0.0001$). b — There were two cases of mild centrilobular necrosis in the control group. c — NOAEL = 0.152 mmol/kg/d (38 mg/kg/day) [Fisher Exact Test].

Table 5. Incidence and Severity of Centrilobular Necrosis in CDBM-Treated Mice

	Dose [mmol/kg/d (mg/kg/day)]					
	0	0.304 (63)	0.610 (127)	1.52 (317)	3.04 (633)	6.10 (1271)
Severity	Percentage of animals					
No cenlob	100%	81%	6%	0	0	
mild	0	19%	28%	0	0	
mild-mod	0	0	16%	0	0	
moderate	0	0	50%	86%	60%	
mod-severe	0	0	0	0	0	
severe	0	0	0	14%	40%	
Total[a,b]	0	19%[c]	94%	100%	100%	
N	13	16	18	14	5	0

a — There was a significant dose-response ($p<0.0001$).
b — There was increasing severity of centrilobular necrosis with dose.
c — NOAEL = 0.152 mmol/kg/d (32 mg/kg/day) [Fisher Exact Test]. There were three cases of mild centrilobular necrosis at the NOAEL.

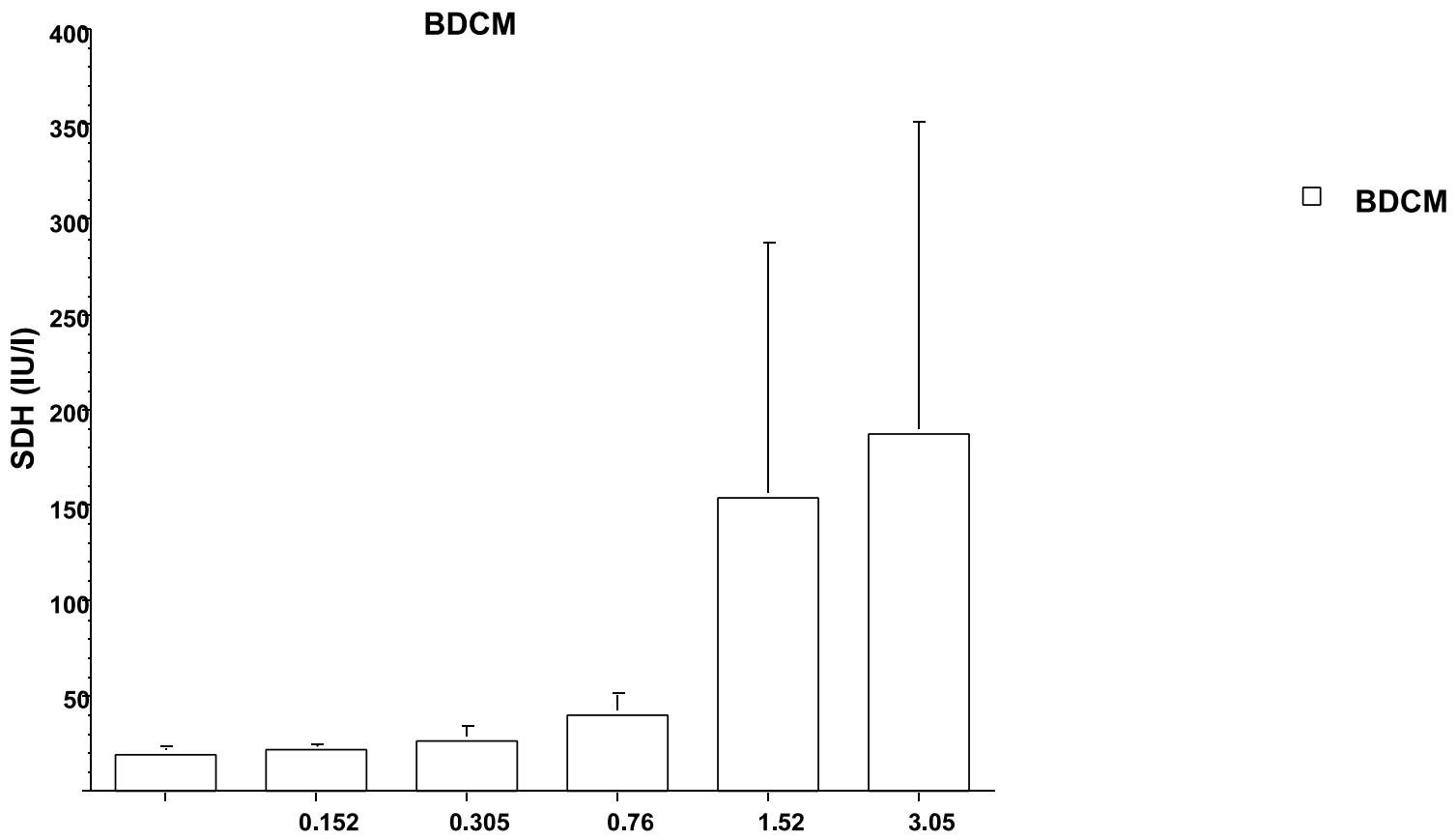


Figure 2
Dose-Related Changes in Mean SDH Levels for BDCM

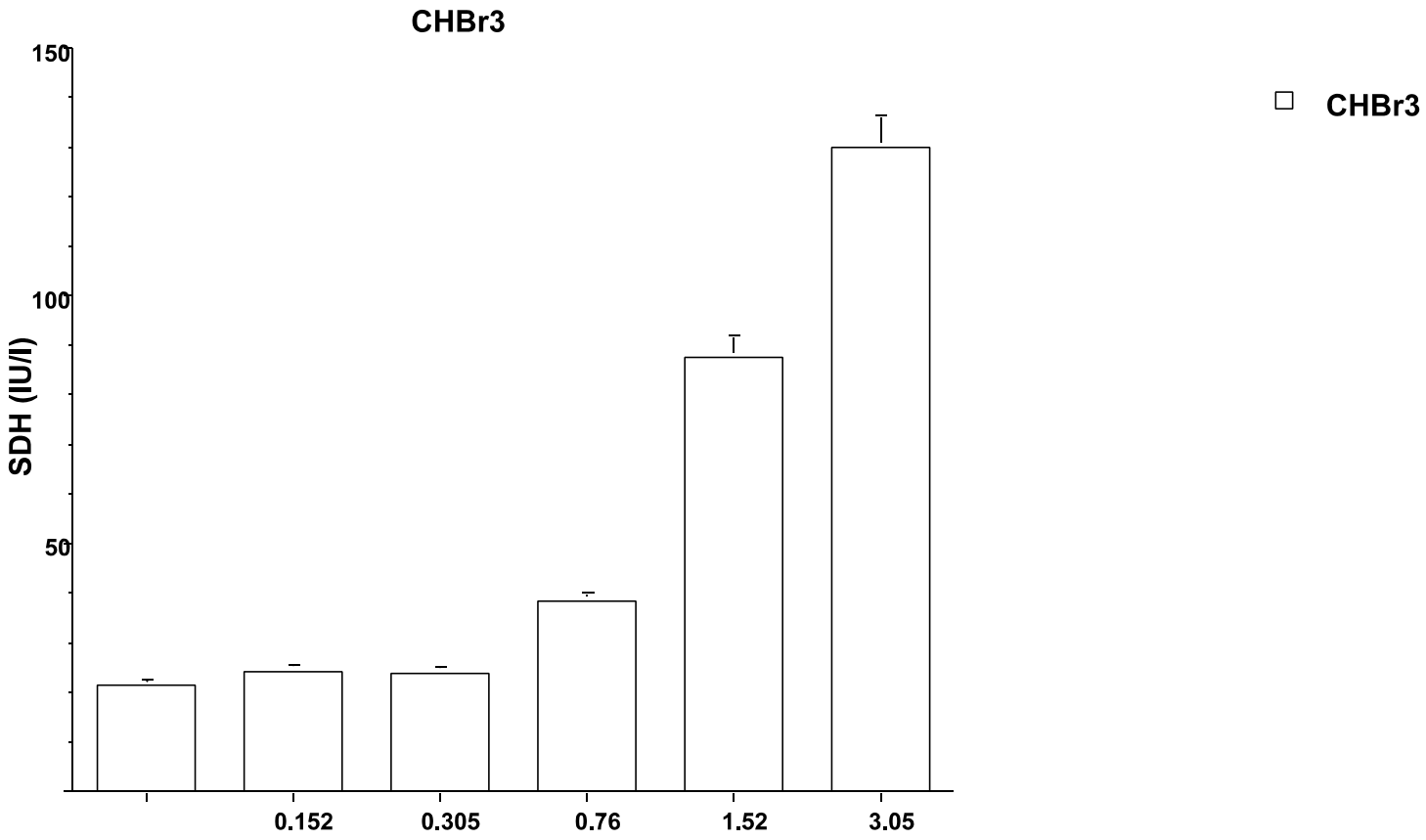


Figure 3
Dose-Related Changes in Mean SDH Levels for $CHBr_3$

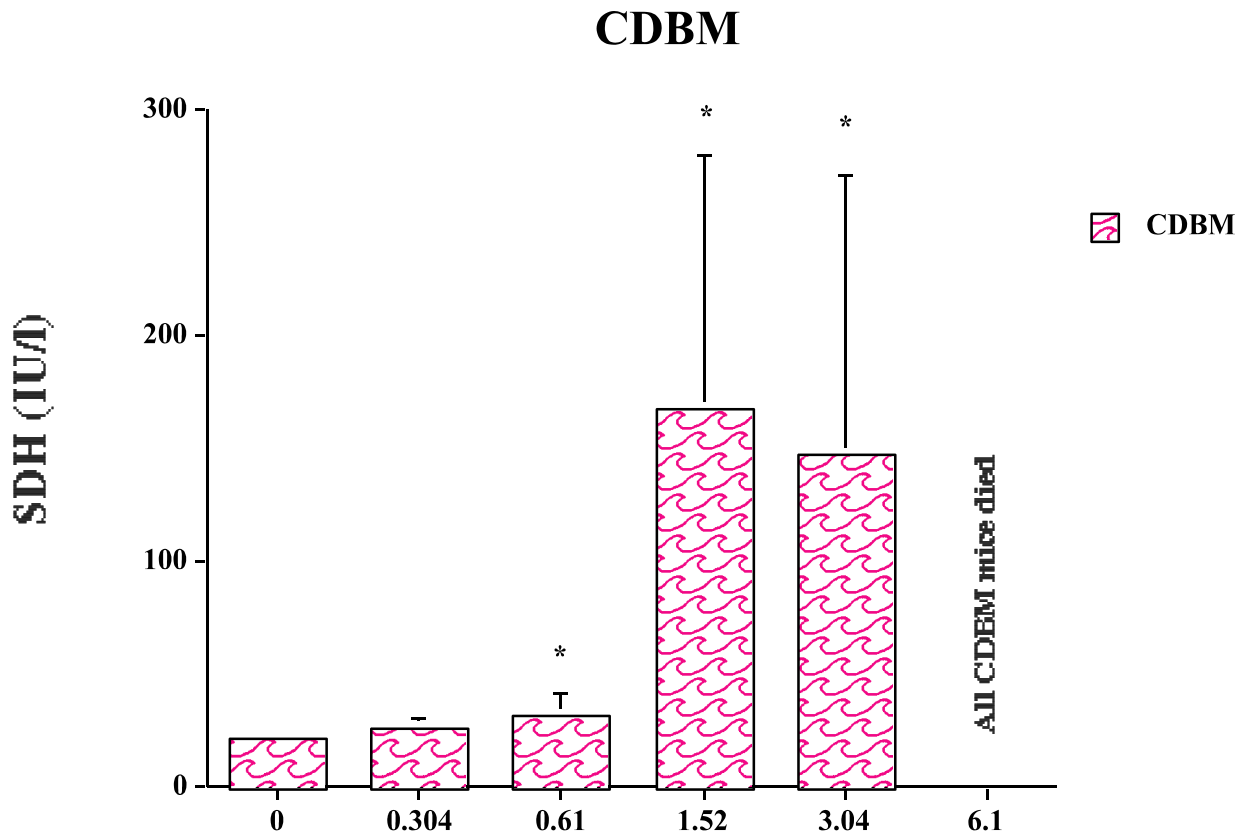


Figure 4
Dose-Related Changes in Mean SDH Levels for CDBM

Mixture Results

The threshold dose-additivity model was fit to the SDH and ALT data because, as Table 1 illustrates, these two endpoints were the most sensitive. Tables 6 and 7 show the observed mean SDH and ALT responses, respectively, for each of the mixtures and for the concurrent single chemical data points. For this analysis to be valid, the observed responses for the concurrent single chemicals should fall within their respective prediction intervals. This is important because the single-chemical experimental results used to build the threshold additivity model were generated by the same laboratory, but earlier than the mixtures and concurrent single chemical data. Thus, the concurrent single chemical data provide a check as to whether it is appropriate to compare the response data from the mixtures experiment with the threshold additivity model predictions. Tables 6 and 7 show that most of the concurrent single chemical data points do fall within or close to their respective prediction intervals, so it is reasonable to evaluate the joint action of the mixtures using this approach.

For the mixtures data, the null hypothesis is that the responses are the result of dose-addition at these particular mixture points. Because the experimental (observed) responses fall within the prediction intervals (expected responses), departure from additivity cannot be claimed. Tables 6 and 7 illustrate that the observed laboratory SDH responses for the two mixtures each fall within their prediction intervals, showing no departures from additivity. The remarkable closeness of the predicted mixture responses to the observed mixture responses for both SDH and ALT provides evidence that the four chemicals are dose-additive for hepatotoxicity in the female CD-1 mouse at the mixture combinations tested.

Table 8 shows the pathology results for the Krasner mixtures and for the concurrent single chemical data. There is a dose-related increase in incidence and severity of centrilobular necrosis as the mixture total dose doubles from 0.436 to 0.872 mmol/kg/d. These data are consistent with the data in Tables 6 and 7, where the SDH and ALT levels also increase as the mixture doses increase.

Discussion

The method of dose addition, with appropriate caveats, has been recommended for assessment of additivity for noncancer health effects (U.S. EPA, 1986). The ability to use isobolographic methods to assess dose addition becomes increasingly difficult as the number of chemicals in the mixture increases. The threshold additivity method of Gennings et al. (1997) overcomes this difficulty. Additionally, and importantly, this methodology provides estimates of variability (e.g., 95% prediction interval, standard deviation) for the predicted response of the mixture.

Table 6

Analysis of Observed SDH Data with Krasner Mixtures
Mixture ratio of (0.24:0.10:0.65:0.01)

Dose (mmol/kg/d)	Sample SDH mean +/- SD (sample size)	Predicted Responses +/- SE	95% Prediction Interval
$CHCl_3$ (2.10)	217.1 +/- 178.2 (5)	58.1 +/- 6.77	[0, 215.5]
$CHBr_3$ (1.00)	56.5 +/- 26.7 (5)	46.4 +/- 3.2	[22.0, 70.7]
BDCM (1.52)	156.9 +/- 149.9 (8)	96.5 +/- 6.74	[0, 201.6]
CDBM (1.20)	116.3 +/- 118.2 (7)	81.0 +/- 5.22	[0, 169.5]
Mixture (Total Dose = 0.436)	33.8 +/- 13.2 (15)	28.9 +/- 4.01	[21.0, 36.8]
Mixture (Total Dose = 0.872)	43.9 +/- 15.8 (19)	40.5 +/- 2.46	[31.9, 49.1]

Table 7

Analysis of Observed ALT Data with Krasner Mixtures
Mixture ratio of (0.24:0.10:0.65:0.01)

Dose (mmol/kg/d)	Sample ALT mean +/- SD (sample size)	Predicted Responses +/- SE	95% Prediction Interval
$CHCl_3$ (2.09)	296.2 +/- 194.2 (5)	91.8 +/- 10.0	[0, 264]
$CHBr_3$ (0.989)	122.4 +/- 138.3 (5)	75.2 +/- 5.98	[0, 198]
BDCM (1.52)	226.0 +/- 160.3 (8)	161.1 +/- 12.4	[46.9, 275]
CDBM (1.20)	203.6 +/- 121.5 (7)	100.0 +/- 10.0	[7.95, 192]
Mixture (Total Dose = 0.436)	37.5 +/- 35.9 (15)	27.4 +/- 3.23	[8.09, 46.7]
Mixture (Total Dose = 0.872)	49.9 +/- 29.3 (19)	48.1 +/- 4.0	[32.7, 63.5]

Table 8. Incidence and Severity of Centrilobular Necrosis in the Krasner Mixtures and Concurrent Single Chemical -Treated Mice

	Dose (mmol/kg/d)						
	Control (0)	Mixture (0.436)	Mixture (0.872)	$CHCl_3$ (2.1)	$CHBr_3$ (1.0)	BDCM (1.52)	CDBM (1.2)
Severity	Percentage of Animals (Incidence)						
No centrilobular necrosis	100% (6)	87% (13)	47% (9)	0	40% (2)[a]	12% (1)	43% (3)
mild	0	13% (2)	37% (7)	20% (1)	40% (2)	0	14% (1)
mild-moderate	0	0	0	60% (3)	20% (1)	0	14% (1)
moderate	0	0	16% (3)	0	0	50% (4)	29% (2)
moderate-severe	0	0	0	0	0	26% (2)	0
severe	0	0	0	20% (1)	0	12% (1)	0
Total centrilobular necrosis	0% (0)	13% (2)	53% (10)	100% (5)	60% (3)	88% (7)	57% (4)
N	6	15	19	5	5	8	7

a — There was one case of severe hepatitis with no centrilobular necrosis in the bromoform group.

The difficulty of full-factorial designs greatly increases as the number of chemicals increases. The threshold addivity model reduces much of this difficulty. The example mixture in this study contained only four chemicals, but the described approach may be applied to mixtures containing many chemicals. The experimental design employed in this study requires substantially fewer animals and dose groups than are required for traditional full-factorial designs, reducing the cost, laboratory resources, and time needed to determine the nature of the interaction of multiple chemicals. This approach is advantageous, particularly when mixtures of certain proportions and dose levels of the individual chemicals are of interest.

The requirement for dose-response curves of the single chemicals is both a potential disadvantage and advantage. These may not be available in the literature. If single-chemical dose-response curves are available, the predicted mixture response can be generated from them, further reducing the time required, the cost, and the number of animals needed — a decided advantage. For this approach to work, the single-chemical dose-response curves must be reproducible in the laboratory where the mixture is tested.

For the mixture of the four THMs, deviation from additivity was not detected for SDH and ALT at the particular mixture composition and concentrations tested. Dose-response trends indicated by the serum enzyme data were consistent with relative liver weights and histopathology results. For these measures of toxicity, the closeness of the predicted mixture responses and the observed mixture response indicates that the four chemicals are dose-additive for hepatotoxicity at the mixture combination tested. Under conditions of dose additivity, the health risks associated with multiple chemical exposures may be predicted with confidence from the single-chemical data.

References

Berenbaum, MC. 1985. The Expected Effect of a Combination of Agents: The General Solution. *J. Theoret. Biol*. 114:413-431.

Berenbaum, MC. 1989. What is Synergy? *Pharmacol. Rev.* 41:93-141.

Gennings, C, Carter, WH Jr. 1995. Utilizing Concentration-Response Data from Individual Components to Detect Statistically Significant Departures from Additivity in Chemical Mixtures. *Biometrics.* 51:1264-1277.

Gennings, C, Schwartz, P, Carter, WH Jr., Simmons, JE. 1997. Detection of Departures from Additivity in Mixtures of Many Chemicals with a Threshold Model. 2:198-211. (Erratum: *J. Agric. Biol. Environ. Stat.* 5(2), 2000)

Krasner, SW, McGuire, MJ, Jacangelo, JG, Patania, NL, Reagan, KM, and Aieta, EM. 1989. The Occurrence of Disinfection Byproducts in U.S. Drinking Water. *J. Am. Water Works Assoc.* 81:41-53.

Munson, AE , Sain, LE, Sanders, VM, Kauffmann, BM et al. 1982. Toxicology of organic drinking water contaminants trichloromethane, bromodichloromethane, dibromochloromethane and tribromomethane. *Environ. Health Perspect.* 46:117-126.

U.S. EPA. 1986. Guidelines for the health risk assessment of chemical mixtures. Federal Register. 51(185):34014-34025.

U.S. EPA. 1999. Integrated Risk Information System (IRIS). Online. National Center for Environmental Assessment, Washington, DC.

PROGRESS REPORT: THE EFFECT OF MIXING RATIO ON MIXTURES OF TRIHALOMETHANE DISINFECTION BY-PRODUCTS

A. McDonald[1], C. Gennings[2], W.R. Hartley[3], L.K. Teuschler[4], A. Thiyagarajah[3], Y.M. Sey[1], S.W. Krasner[5], J.E. Simmons[1]

[1]NHEERL, U.S. EPA, RTP, NC
[2]Virginia Commonwealth University, Richmond, VA,
[3]Tulane University, Environmental Diseases Prevention Research Center, New Orleans, LA,
[4]NCEA, U.S. EPA, Cincinnati, OH,
[5]Metropolitan Water District of Southern California, La Verne, CA

August 1999

Abstract

Chlorination of water containing organic matter results in the production of a variety of halogenated by-products, including the trihalomethanes (THMs) chloroform ($CHCl_3$), bromodichloromethane (BDCM), chlorodibromomethane (CDBM), and bromoform ($CHBr_3$). Water treatment processes that involve ozonation/chloramination decrease the total amount of THMs formed but alter their relative distribution. Thus, the objective of this investigation was to determine the effect of mixing ratio on mixtures of the four THMs. Female CD-1 mice (~65 to 70 days old) were gavaged daily for 14 days with the THM mixtures in 10% Alkamus® El-620. Hepatotoxicity was assessed on day 15 by serum sorbitol dehydrogenase (SDH) and by histopathology. Relevant mixing ratios were selected from a study in which a high bromide water was treated either by prechlorination/post-chlorination (CL) or by pre-ozonation/post-chloramination (OZCM). The proportions of the CL mixture were $CHC1_3$, 0.319; $CHBr_3$, 0.049; BDCM, 0.342; and, CDBM, 0.290 while the proportions of the OZCM mixture were $CHC1_3$, 0.655; $CHBr_3$, 0.029; BDCM, 0.187; and, CDBM, 0.129. Three experiments were conducted, each testing a different total dosage of these two mixtures: 0.05, 1.50, and 3.0 mmol/kg/day. At each total mixture dosage, the SDH values for the CL and OZCM mixtures did not differ significantly. Similarly, the proportion of mice with centrilobular hepatocellular necrosis did not differ significantly between the two mixing ratios. In conclusion, at the same dosage, mixing ratio, or at least the two mixing ratios tested here, had little/no apparent effect on the hepatotoxicity of THM mixtures. Data such as these may help in selecting optimal water treatment strategies. As the total mixture dose was held constant in each experiment, differences resulting from the amount of THMs produced by each treatment are not reflected in these experiments but are worthy of future investigation.

Introduction

Chlorination of drinking water has resulted in dramatic decreases in outbreaks of water-borne diseases and is rightly considered one of the public health triumphs of the 20th century. However, disinfection by chlorination of water containing organic matter results in a vast array of disinfection by-products (DBPs). Although the majority, by mass, of DBPs remain to be identified, the known DBPs include a number of toxic and carcinogenic chemicals, such as chloroform ($CHCl_3$) bromodichloromethane (BDCM), the chlorinated furanone, MX, and the haloacetic acids. By mass, typically the most prominent chemical class of the DBPs is the trihalomethanes (THMs). The four THMs present in disinfected water are $CHCl_3$, BDCM, chlorodibromomethane (CDBM), and bromoform ($CHBr_3$).

The amounts and relative proportions of the DBPs, including the THMs, formed during chlorination are a function of the amount of chlorine used as well as a variety of characteristics of the water undergoing treatment, including its humic acid concentration, bromide ion content, pH, and temperature. In an effort to decrease the amount of THMs formed, the use of alternative water treatment processes is increasing. Ozone is an effective treatment for disinfection of water. However, ozone as a primary treatment requires a secondary treatment such as chlorine or chloramine. The amounts and relative proportions of the DBPs formed are a function of the water quality factors listed above as well as the concentration of both ozone and the post-treatment chemical, their contact times, and the time between the two treatments.

Treatment of water with ozonation/chloramination decreases the total amount of THMs formed and alters the relative distribution of individual components. Because relative proportion of chemicals in a mixture has the potential to influence the toxicity of the mixture, the toxicity of mixtures of DBPs from water disinfected by different treatments may differ not only because of differing amounts of total DBPs but also because of differing relative proportions of the DBPs comprising the mixture.

Objective

The objective of this study was to examine the effect of mixing ratio on the hepatotoxicity of mixtures of $CHCl_3$, BDCM, CDBM, and $CHBr_3$ (the four THMs formed during chemical disinfection of drinking water). A further objective was to use environmentally relevant mixing ratios. The mixing ratios used were based on the relative proportions of the THMs formed during either pre- and post-treatment with chlorine (CL mixture) or pre-treatment with ozone followed by post-treatment with chloramine (OZCM mixture) (Krasner et al., 1989).

Methods

Female CD-1 mice (Charles River Breeding Laboratory, Raleigh, NC), 65 to 70 days of age at the beginning of the experiments, were used. Groups of 8 to 20 mice were exposed daily by oral gavage for 14 days to mixtures of the four THMs in an aqueous-based vehicle containing 10% Alkamuls® EL-620. The gavage volume was 10 mL/kg. Animals were gavaged in the

morning between 8 a.m. and noon with gavage solutions made fresh daily in gas-tight vials. Three experiments were conducted. Each experiment included a total mixture dosage that was the same for CL and OZCM (see below), a control group, and a single chemical dose group for each THM in the mixture ($CHCl_3$, $CHBr_3$ and BDCM at 1.52 mmol/kg/day and CDBM at 0.76 mmol/kg/day).

On the morning following the 14th day of dosing, anoxia was induced with carbon dioxide. Mice were bled from the abdominal aorta. Serum was prepared and stored at -80°C until analyzed for the activities of sorbitol dehydrogenase (SDH), alanine aminotransferase (ALT), and aspartate aminotransferase (AST) by automated procedures (Cobas Fara II) with appropriate reagents and standards.

Slices from the left hepatic lobe were preserved in 10% phosphate-buffered formalin. Serial sections, 5 μm thick, were cut from paraffin-embedded tissue. Slides were stained with hematoxylin and eosin and examined blind (i.e., without knowledge of the dose group) by bright-field microscopy. Centrilobular necrosis was characterized as none (no necrotic hepatocytes seen), mild (less than 10% of the hepatocytes involved), mild to moderate (less than 50% of the hepatocytes involved), moderate (approximately 50% of the hepatocytes involved), moderate to severe (more than 50% of the hepatocytes involved), and severe (more than 75% of the hepatocytes involved).

Chlorination and Ozonation Mixtures

Table 1 shows the relevant mixing ratios (on a molar basis) selected from a study in which a high bromide water (150 μg/L) was treated by either pre-chlorination / post-chlorination (CL) or by pre-ozonation / post-chloramination (OZCM) (personal communication from S. Krasner, based on data shown in Krasner et al., 1989) . It is important to note that the mixing ratios for the two treatment conditions were selected from a study in which the same input water stream was split and underwent two different disinfection scenarios. CL treatment formed 101 μg/L (0.61 μmol/L) total THMs and OZCM treatment formed 2 μg/L (0.014 μmol/L) total THMs. In this study, CL treatment formed much higher concentrations of total THMs than did OZCM.

Table 1
Mixing Ratios for the CL and OZCM Treatment Processes

THM	CL Ratios		OZCM Ratios
$CHCl_3$	0.319		0.655
$CHBr_3$	0.049		0.029
BDCM	0.342	0.187	
CDBM	0.290	0.129	

Three experiments were conducted, each testing a different total dosage of the two mixtures: 0.05, 1.50, and 3.0 mmol/kg/day. The high-dose CL mixture and the low-dose OZCM mixture approximately match the relative difference in total THMs from the water treatment study described above.

Data Analysis Methods

Serum data

The mean responses to the CL and OZCM mixtures were compared across mixture dose levels for SDH, ALT, and AST using analysis of variance (ANOVA). The model was parameterized to include terms for group (CL vs OZCM), mixture dose level (0.05, 1.50, and 3.0 mmol/kg/day), and group by dose interactions. Homogeneity of variance was assessed by Hartley's test. The Shapiro-Wilks test was used to examine normality in the residuals. Log transformations of the SDH, ALT, and AST data were needed to satisfy the assumption of normality. The criterion of significance was 0.05.

Histopathology data

Histopathology data from the CL and OZCM mixture experiments were analyzed for the presence or absence of hepatocellular necrosis. The proportion of animals with necrosis was analyzed by logistic regression using an analysis of variance parameterization to determine whether differences existed between CL and OZCM at the 0.05 level of significance. Likelihood ratio tests were used to test for dose and group effects.

Results

Serum Data

Figures 1, 2, and 3 illustrate there were no differences in the SDH, ALT, and AST serum enzyme levels between the CL and OZCM mixing ratios for the total mixture dosages of 0.05, 1.5, and 3.0 mmol/kg/day, respectively. The interaction terms (group x dose) that test for whether differences in responses between CL and OZCM are affected by dose were not statistically significant. Statistically significant differences were not detected between CL and OZCM at any dose level for any endpoint. For both CL and OZCM individually however, there was a significant dose effect at 0.05, 1.50, and 3.0 mmol/kg/day, indicating an increasing effect with increasing dose.

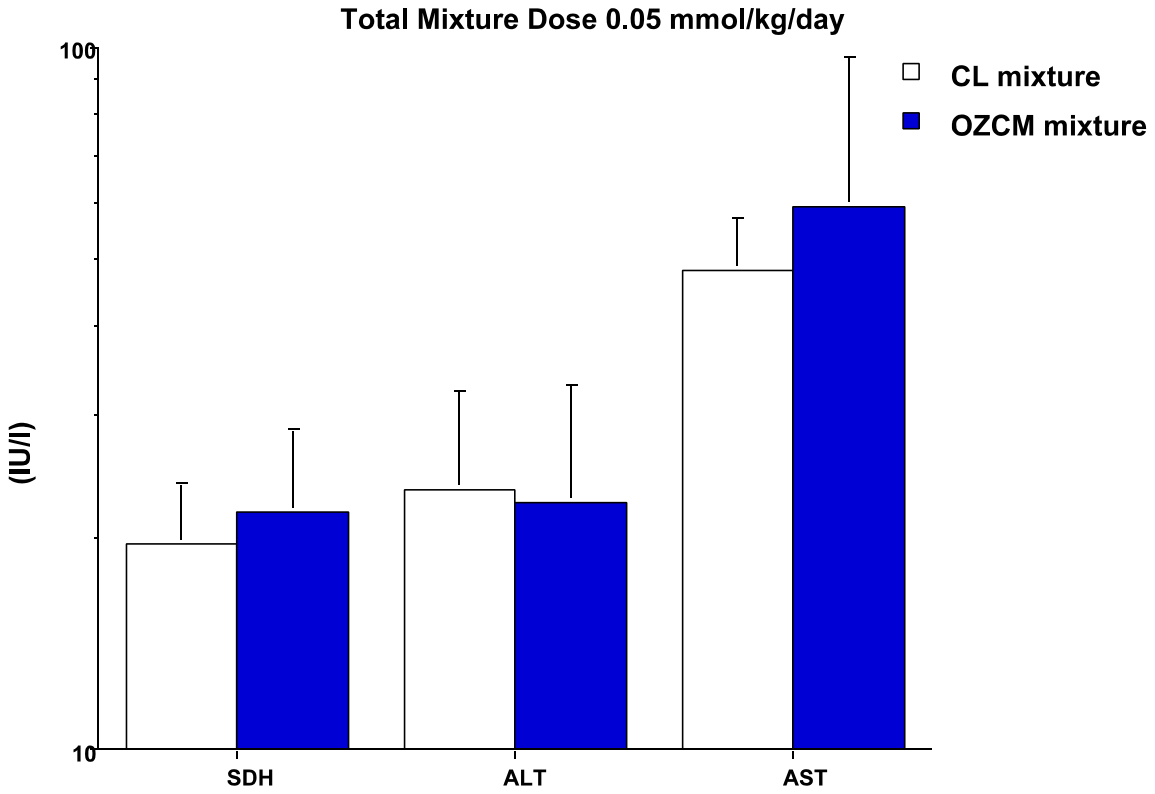


Figure 1
Serum Enzyme Levels for CL and OZCM Mixtures in CD-1 Mice

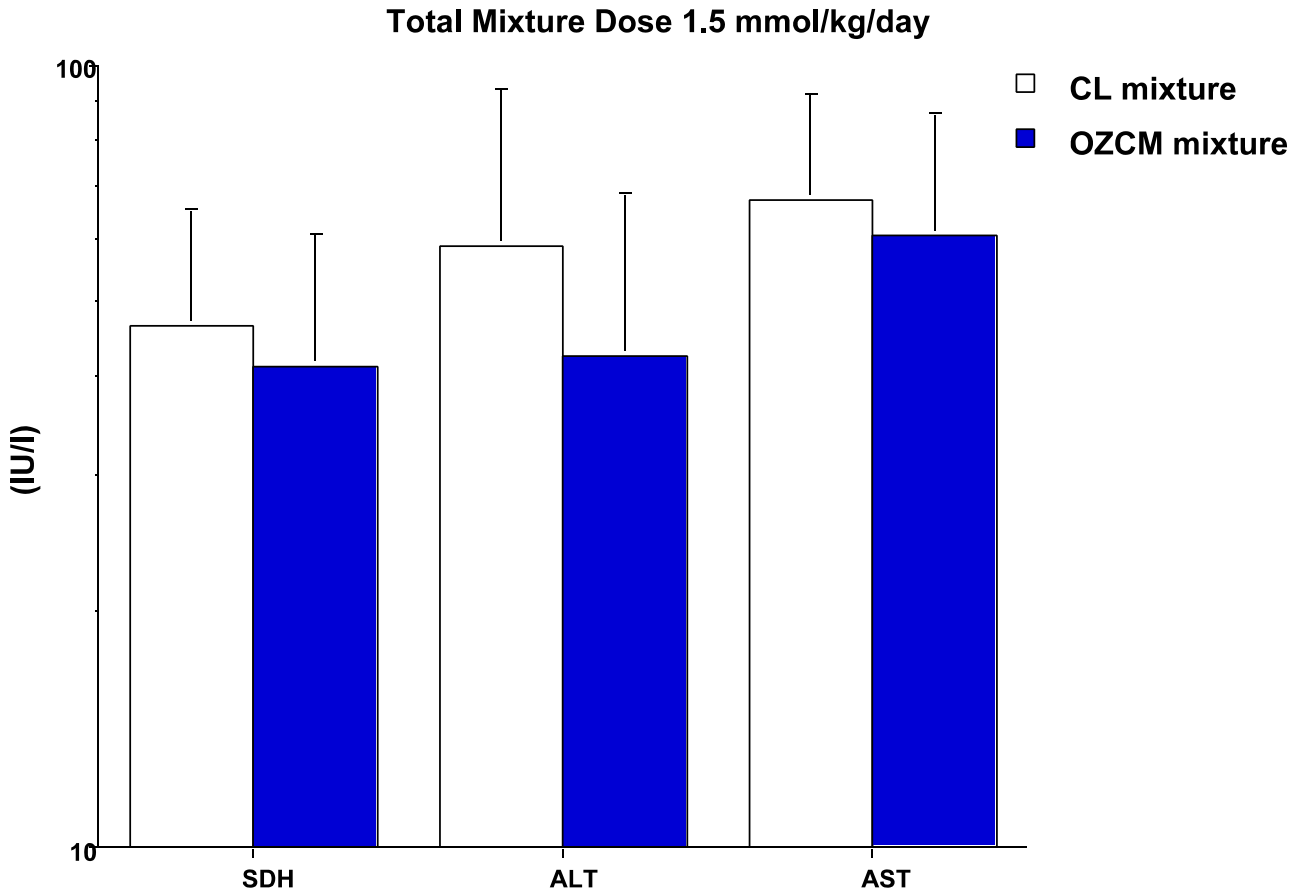


Figure 2
Serum Enzyme Levels for CL and OZCM Mixtures in CD-1 Mice

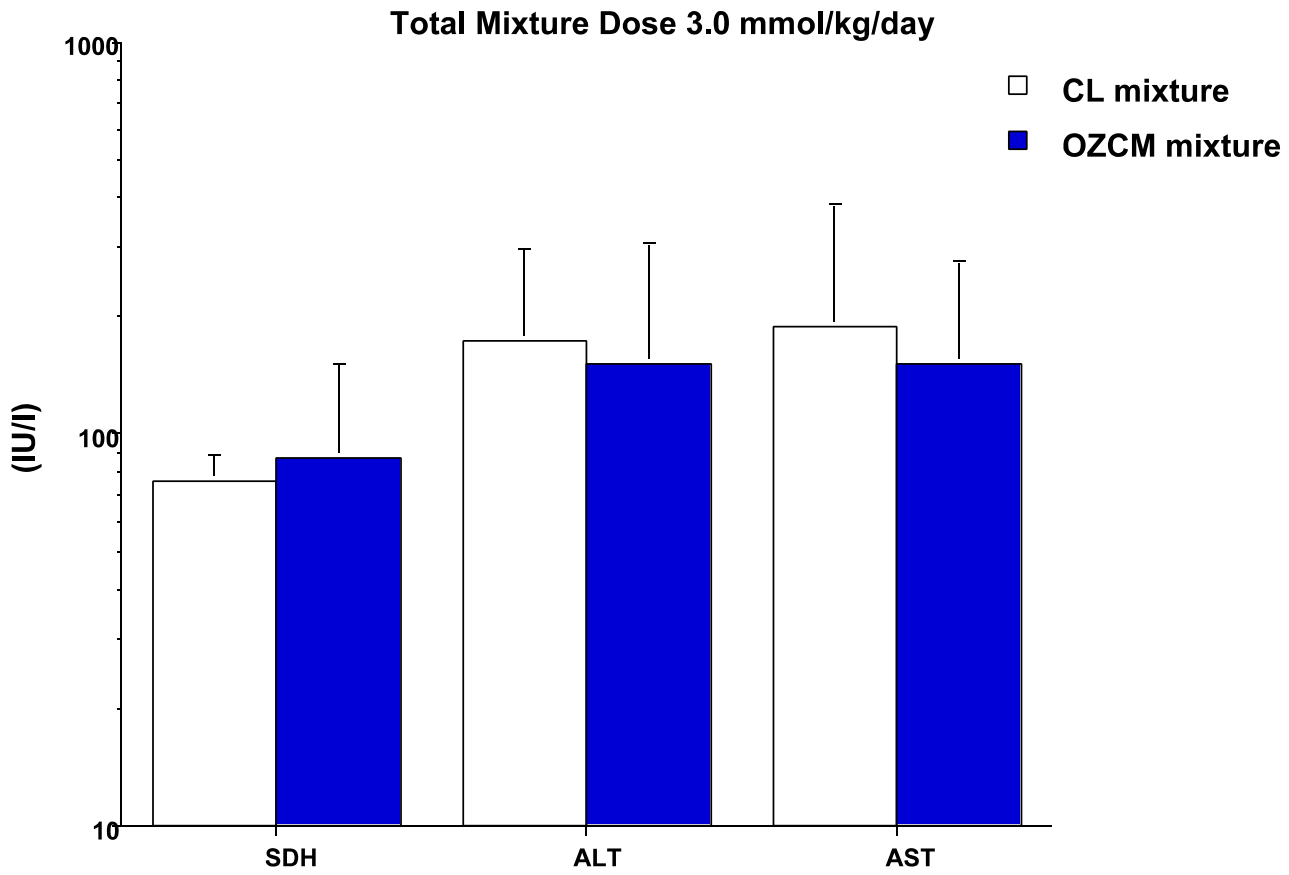


Figure 3
Serum Enzyme Levels for CL and OZCM Mixtures in CD-1 Mice

Histopathology Data

Table 2 shows the incidence and severity of centrilobular hepatocellular necrosis increased with total mixture dosage for both CL and OZCM. At 0.05 mmol/kg/day of either CL or OZCM, centrilobular necrosis (CN) was not observed. At 1.5 mmol/kg/day, ~ 65% of CL mice showed CN that varied in severity from mild to moderate-severe, and ~ 44% of OZCM mice had necrosis that ranged from mild to moderate. At 3.0 mmol/kg/day, all CL mice had CN that was predominately moderate to moderate-severe. At this dosage (3.0 mmol/kg/day), the vast majority (~94%) of OZCM mice had CN ranging from mild to severe.

The proportion of mice with hepatocellular necrosis did not differ between the CL and OZCM mixtures at either 0.05, 1.50 or 3.0 mmol/kg/day. However, as total dose increased, there was a statistically significant increase in the proportion of mice with hepatocellular necrosis.

Discussion

This series of experiments was designed to examine the effect of mixing ratio on the hepatotoxicity of the four THMs formed during disinfection of drinking water by either chlorination (pre- and post-treatment, CL) or pre-treatment with ozone followed by post-treatment with chloramine (OZCM). An advantage of the present work is that the mixing ratios selected to represent the two different disinfection scenarios came from chemical analysis of a source water split into two streams, with each stream undergoing either CL or OZCM. As source water characteristics are known to have a strong impact on the DBPs formed, use of the same source water, while not removing the influence of source water on toxicity, ensured that it was the same for the two disinfection schemes considered here. Additional studies where the same source water undergoes different treatment scenarios are needed to better characterize the interaction between source water and disinfection scheme on DBP production.

The data presented here provide evidence that mixing ratio, at least the two mixing ratios tested here, had little or no apparent effect on the hepatotoxicity of THM mixtures. For SDH, ALT, and AST, as well as centrilobular necrosis, statistically significant differences were not detected that could be attributed to mixing ratio. However, it must be noted that mixing ratio (i.e., the relative proportion of the THMs) is not the only DBP factor affected by changing from CL to OZCM. For example, the amount of THMs produced is decreased significantly by OZCM. In addition, because OZCM produces other DBPs not produced by CL (e.g., bromate, cyanogen, chloride); thus, other mixtures issues must also be considered.

Our future directions include examining the effect of the relative differences in total THMs formed by CL and OZCM. We are also conducting experiments for each of the six possible binary combinations of THMs. These data will be used to validate and/or assess three different statistical and risk assessment methods: the threshold additivity model,

Table 2

Incidence and Severity of Centrilobular Hepatocellular Necrosis

Mixture Dosage (mmol/kg/day)	0.05		1.50		3.0	
	CL	OZCM	CL	OZCM	CL	OZCM
N	17	19	17	18	16	16
Lesions						
None	100%	100%	35.3%	55.6%	0%	6.3%
Mild	0%	0%	23.5%	22.2%	12.5%	12.5%
Mild-moderate	0%	0%	5.9%	16.7%	18.9%	25.0%
Moderate	0%	0%	17.6%	5.6%	31.3%	31.3%
Moderate-severe	0%	0%	17.6%	0%	37.5%	12.5%
Severe	0%	0%	0%	0%	0%	12.5%
Total % Affected	0%	0%	64.7%	44.4%	100%	93.8%

proportional-response addition, and the interactions-based hazard index. We also plan to extend this research beyond THMs to other chemical classes of DBPs.

References

Krasner, S.W., M.J. McGuire, J.G. Jacangelo, N.L. Patania, K.M. Reagan and E.M. Aieta. 1989. The occurrence of disinfection by-products in U.S. drinking water. JAWWA. 81: 41-53.

CATEGORICAL REGRESSION ANALYSIS OF BROMODICHLOROMETHANE LIVER TOXICITY AND PATHOLOGY DATA

LK Teuschler[1], WR Hartley[2], A Thiyagarajah[2], JC Lipscomb[1], JE Simmons[3]

[1]NCEA, U.S. EPA, Cincinnati, OH
[2]Tulane University Medical Center, New Orleans, LA
[3]NHEERL, U.S. EPA, RTP, NC

September 1999

Abstract

A categorical regression (CR) procedure is being investigated as a way of comparing new toxicity data with existing benchmarks such as a Reference Dose (RfD), and as a method for predicting pathology scores from other hepatotoxic endpoints. CR was used to model data on liver toxicity and histopathology scores from female CD-1 mice exposed for 14 days by gavage to doses of 0, 0.152, 0.305, 0.76, 1.52, and 3.05 mmol/kg/day bromodichloromethane (BDCM), a common by-product of drinking water disinfection processes. The toxic effects included both continuous endpoints (liver enzyme elevations, liver weight changes) and categorical data (centrilobular necrosis severity designations). The centrilobular necrosis data were classified into ordered categories of severity (i.e., category 1 = none detected, category 2 = mild, mild-moderate to moderate, and category 3 = moderate-severe to severe) at the individual animal level. Two different CR models were then constructed. The *first CR model* was used to predict the probability of centrilobular necrosis at the Lowest-Observed-Adverse-Effect Level (LOAEL) used to estimate the BDCM RfD. At the animal chronic LOAEL of 17.9 mg/kg/day based on renal cytomegaly in male mice, which was used to calculate the RfD, the probability of observing mild centrilobular necrosis or greater was 0.01 (95% CI of 6E-4, 0.14), consistent with the fact that the kidney effect should be a more sensitive endpoint. The *second CR model* was applied to a test data set of eight mice dosed with 1.2 mmol/kg/d BDCM to predict the occurrence and severity of centrilobular necrosis based on information on the liver serum endpoint, alanine aminotransferase (ALT), level and the relative liver weight. These predictions were then compared with the actual pathology scores. For occurrence, seven of the eight observed pathology severity scores matched the prediction of at least mild centrilobular necrosis. For severity, in four of the eight mice, the observed pathology severity score corresponded exactly to the most likely severity, as predicted by the model. The CR procedures demonstrated here are a useful way to evaluate histopathology data that are categorical and qualitative in nature.

Introduction

Histopathology results are generally expressed as the number of experimental animals that fall into one of a series of qualitative categories of response (e.g., mild centrilobular vacuolation, mild centrilobular necrosis, severe centrilobular necrosis) ordered by increasing severity of the effect. Thus, these data may be typically analyzed using less than optimal techniques, such as pairwise significance testing of mean responses, tests for linear trend, or Fisher exact tests, that are unable to take the ordered severity of the effect into account. Categorical regression (CR) is an effective statistical procedure in this circumstance because it regresses ordered categorical data on continuous variables to predict the likelihood that a particular category of severity will be observed. Hertzberg and Miller (1985) and Hertzberg (1989) first proposed the CR procedure as a way to estimate the risk of adverse health effects from chemical exposures above the Reference Dose (RfD); several others illustrated its use with various chemicals (Farland and Dourson, 1992; Rao et al., 1993; Dourson et al., 1997; Guth et al., 1997). The regression procedure can be used for many types of endpoints including quantal and continuous data, quantitative and qualitative data, and data representing various levels of severity. Several explanatory (independent) variables can be incorporated, such as dose, exposure duration, organ weights, or enzyme levels, and the analysis can be stratified by species (Guth et al., 1991, 1997).

In this research, two categorical regression (CR) models were constructed to analyze the liver pathology data from female CD-1 mice exposed to bromodichloromethane (BDCM), a common disinfection by-product found in drinking water. The *first CR model* was constructed to predict the probability of centrilobular necrosis at the Lowest-Observed-Adverse-Effect Level (LOAEL) used to estimate the BDCM Reference Dose (RfD). The RfD is defined as an estimate of a lifetime daily exposure to the human population likely to be without an appreciable risk of an adverse effect (U.S. EPA, 1999). The RfD is based on the most sensitive critical effect that has been observed in humans or found in animal experimental data; thus, given adjustments for doses across species, the LOAELs for other effects, such as hepatotoxicity, are likely to be observed at higher dose levels. The RfD for BDCM was calculated from a chronic animal LOAEL of 17.9 mg/kg/day based on renal cytomegaly in male B6C3F1 mice (an effect considered minimal in the absence of renal function impairment) (NTP, 1986; U.S. EPA, 1999).

The *second CR model* was constructed to predict the occurrence and severity of centrilobular necrosis based on information on the liver serum endpoint, alanine aminotransferase (ALT) level and the relative liver weight. A linear regression analysis was used previously for this purpose for the histologic evaluation of the effects of pretreatment with various organics on chloroform-induced hepatotoxicity in male rats (Plaa and Hewitt, 1989). In that case, the percent of abnormal, necrotic, or degenerated hepatocytes was regressed on ALT and relative liver weight endpoints (among others), resulting in elevations in the logarithm of ALT as the variable most highly correlated with the histopathologic alterations. Relative liver weight was only weakly correlated. The advantages of using a CR approach over linear regression is that the severity categories generally used to characterize histopathology can be accommodated directly by the method and not converted into continuous, normally distributed

data (e.g., percentages). Furthermore, predictions can be made of not only the probability of occurrence, but also the probability of observing a particular level of severity for the endpoint of concern, i.e., in the BDCM case, centrilobular necrosis.

Methods

Experimental Methods and Data Collection

Female CD-1 mice (N=8 to 20), 65 to 70 days old, were exposed by oral gavage for 14 days to BDCM. Metabolism cages (2/cage) were used to house the animals. The experiment was performed in two blocks, distributing part of the animals from each dose group into each block. BDCM was given in an aqueous vehicle (10% Alkamulus EL-620) at a constant volume of 10 ml/kg at the following dose levels: 0, 0.152, 0.305, 0.76, 1.52, or 3.05 mmol/kg/d. Animals were dosed daily using freshly prepared solutions. On Day 15, the mice were sacrificed. Several serum endpoints and body and liver weights were evaluated at sacrifice. The percent relative liver weight (Pcliv) was calculated as the liver to body weight ratio at the time of sacrifice times one hundred. Serum was analyzed for indicators of hepatic toxicity: sorbitol dehydrogenase (SDH), alanine aminotransferase (ALT), and aspartate aminotransferase (AST). Liver tissues were preserved for histopathologic evaluation in 10% phosphate buffered formalin. Serial sections, 5 μm thick, were cut from paraffin-embedded tissue. Two of six slides of the liver were stained with hematoxylin and eosin and examined blind (without knowledge of dose group) by brightfield microscopy.

Statistical Analysis

Categories of pathological staging were classified into three severity groups representing centrilobular necrosis (i.e., category 1 = none detected, category 2 = mild, mild-moderate to moderate, and category 3 = moderate-severe to severe). An example of individual BDCM mouse data with these severity designations is shown in Table 1. For this application, it is assumed that the severity of centrilobular necrosis is related to the explanatory variables (dose in *model 1*; ln(ALT) and PcLiv in *model 2*) by using a cumulative logistic function:

$$\ln(\pi_i/(1-\pi_i)) = \alpha_i + \beta_1 X_1 + \beta_2 X_2 + ... \beta_n X_n \qquad (1)$$

where

π_i	=	p(severity $\leq$ category i)
X_i	=	explanatory variable i
β_i	=	parameter estimate for variable X_i
α_i	=	intercept or “cut point” between categories

Table 1

Example Hepatotoxicity Data from Bromodichloromethane Exposure

Mouse ID #	Dose mmol/kg/d	ALT IU/l	AST IU/l	SDH IU/l	% Relative Liver Weight	Pathology Severity Score*	Pathology Results**
8	0	21	64	18	5.12	1	Nothing detected
10	0.152	32	69	26	4.01	1	Mild-mod cenlob vacuolation
66	0.305	32	47	28	3.97	1	Mild cenlob vacuolation
52	0.76	35	66	33	5.27	2	Mod cenlob necrosis Mod cenlob vacuolation Mod zone 2 vacuolation
83	1.52	476	249	403	6.0	2	Mod cenlob necrosis Mild-mod ballooned cell necrosis
81	3.05	277	242	249	7.26	3	Sev cenlob necrosis Mod-sev ballooned cell necrosis

*Pathology score for centrilobular necrosis: 1=none; 2=mild, mild-moderate, moderate; 3= moderate-severe, severe
**Abbreviations: Mild = mild; Mod=Moderate; Sev=Severe; Cenlob=Centrilobular

Solving for $1-\pi_i$, the risk of effects worse than severity (S) level "i" is

$$p(S > i) = 1 / [1+\exp(\alpha_i + \beta_1 X_1 + \beta_2 X_2 + ... \beta_n X_{n)}]. \quad (2)$$

Thus, by adding or subtracting probabilities, it is easy to calculate the probability that severity is equal to a given severity category (e.g., P(S=2) = [P(S>1) - P(S>2)]).

As is shown in Equation (1), π_i is defined as the probability of observing a response of a certain severity i or less. For a model with one explanatory variable, the cumulative logit model in Equation (1) specifies that a series of parallel straight lines with a common slope parameter across severity categories defines the relationship between the cumulative logits [i.e., the $\ln(\pi_i/(1-\pi_i)$ terms] and the response variable. For a model with more than one explanatory variable, these become a series of parallel planes. In addition, these models have a proportional odds assumption (i.e., equal slopes) that can be tested using a chi-square test. This assumes that the log odds ratio for two different vectors of the explanatory variables is proportional to the difference between these vectors with the same proportionality constant for every category.

Overall model significance is also tested, as well as each individual parameter estimate, much like that which is done in a simple linear regression procedure. Thus, there is confidence that the model characterizes the relationships well and that the explanatory variables are useful in describing the severity of centrilobular necrosis that is seen. All calculations were made using the SAS LOGISTIC Procedure (SAS Institute, 1990).

Results

General

Table 2 summarizes the histopathology data for the BDCM-exposed mice. Note the dose-related increases of liver enzyme levels, percent relative liver weight and severity of centrilobular necrosis. The four photomicrographs (Figures 1 to 4) show examples of the increase in severity of centrilobular necrosis with dose of BDCM. Although the summary in Table 2 shows general trends in the data, input to each CR model was done at the individual animal level (see Table 1). As shown in Tables 3 and 4, both *models 1 and 2* met the requirements of the proportional odds assumption, and each model was statistically significant in terms of its overall construction and its parameter estimates.

Model 1

Table 5 shows some predicted probabilities generated from *model 1*, which regressed severity categories on one explanatory variable, dose. Doses are given in both mmol/kg/day and in mg/kg/day. The three columns on the far right of Table 5 list the probabilities of observing a particular severity level of centrilobular necrosis, given a specific dose. The other probability

Table 2

Summary Hepatotoxicity Data for CR Input to *Models 1 and 2*

Mice N	Dose mmol/kg/d	ALT μ (σ) IU/l	AST μ (σ) IU/l	SDH μ (σ) IU/l	% Relative Liver Wgt μ (σ)	Pathology Severity Score*	Number of Responders at Pathology Severity Score
10	0	19.2 (24)	49.1 (14.9)	19.3 (3.8)	4.35 (0.57)	1	10
10	0.152	21.3 (4.9)	51.1 (10.8)	21.4 (3.4)	4.19 (0.38)	1	10
11	0.305	25.8 (6.0)	47.9 (7.7)	26.6 (8.0)	4.65 (0.69)	1 2	10 1
11	0.76	41.9 (17.3)	60.5 (17.0)	39.6 (12.4)	5.2 (0.37)	1 2	2 9
10	1.52	236.7 (143.3)	185 (108.0)	154.4 (133.8)	6.21 (1.0)	2 3	6 4
8	3.05	422.2 (281.2)	329.2 (201.1)	187.4 (163.9)	7.98 (0.39)	3	8

*Pathology score for centrilobular necrosis: 1=none; 2=mild, mild-moderate, moderate; 3= moderate-severe, severe

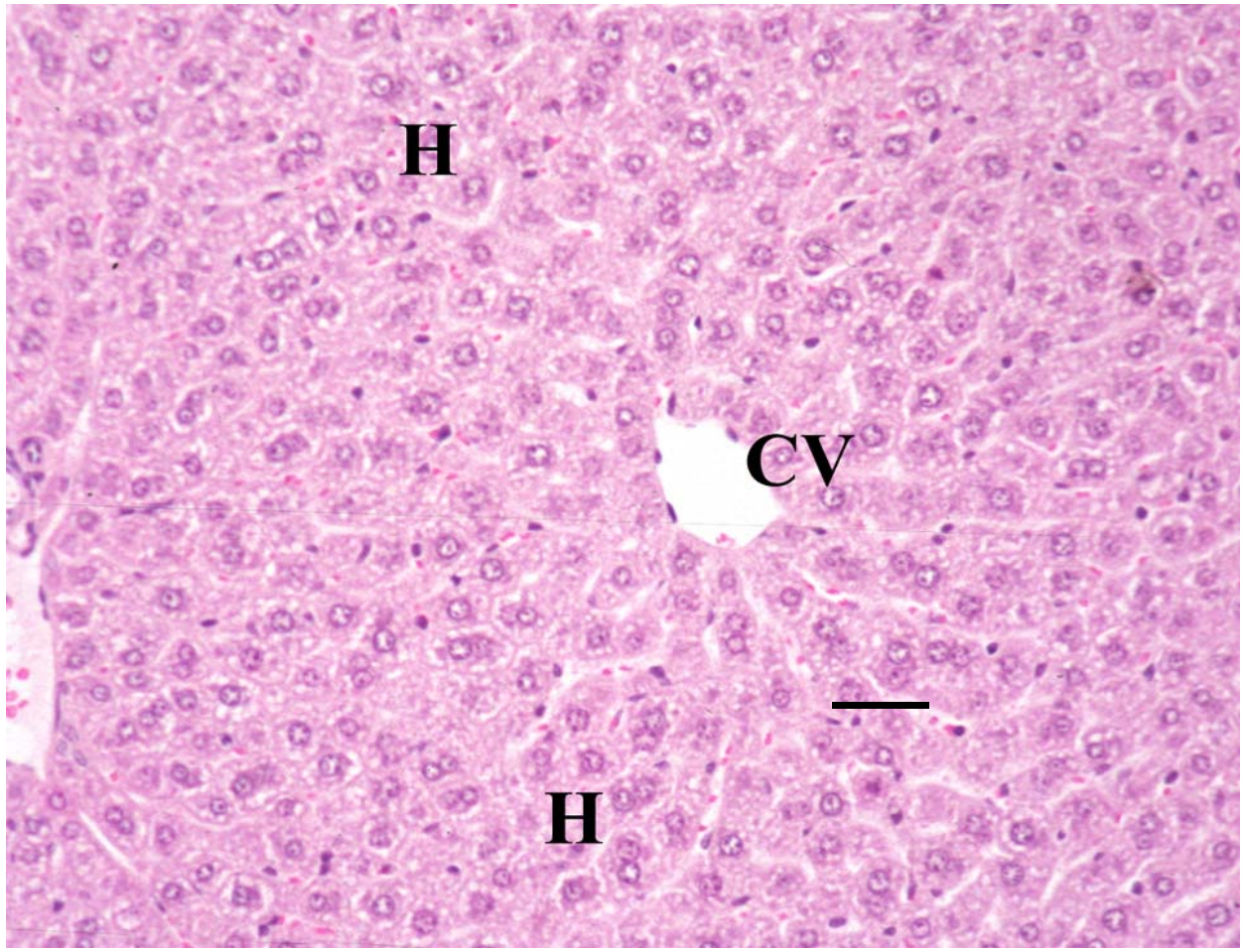


Figure 1. Example of severity category 1. Liver of a control mouse showing no centrilobular necrosis. Hepatocytes (H) and central vein (CV). H&E, bar=30 μm.

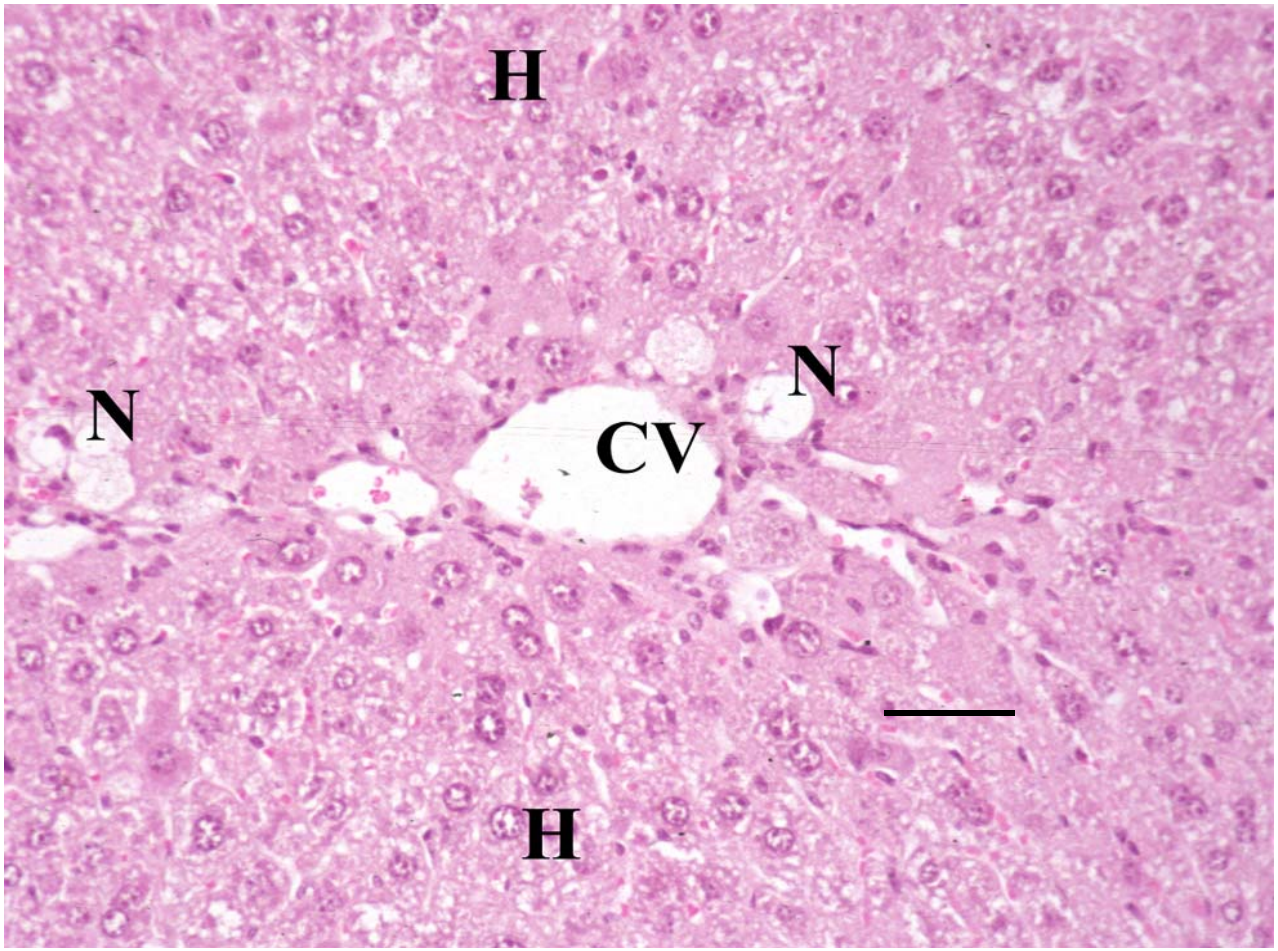


Figure 2. Example of severity category 2. Liver of a mouse treated with BDCM (0.76mmol/Kg) showing a mild centrilobular necrosis (N). Hepatocytes(H) and central vein (CV). H&E, bar = 40 μm.

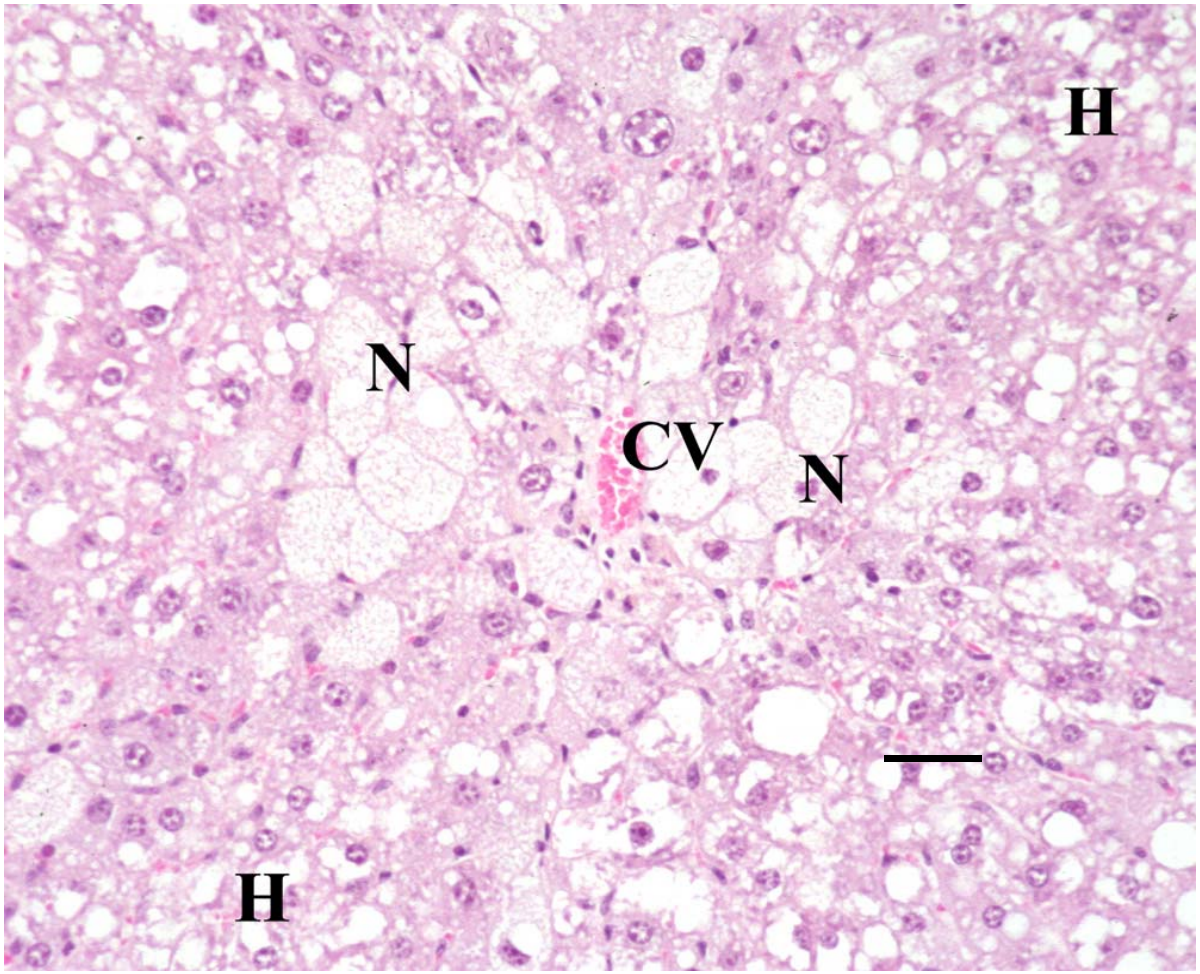


Figure 3. Example of severity category 2. Liver of a mouse treated with BDCM (1.52 mmol/Kg) showing a moderate centrilobular necrosis (N). Hepatocytes (H) and central vein (CV). H&E, bar = 30 μm.

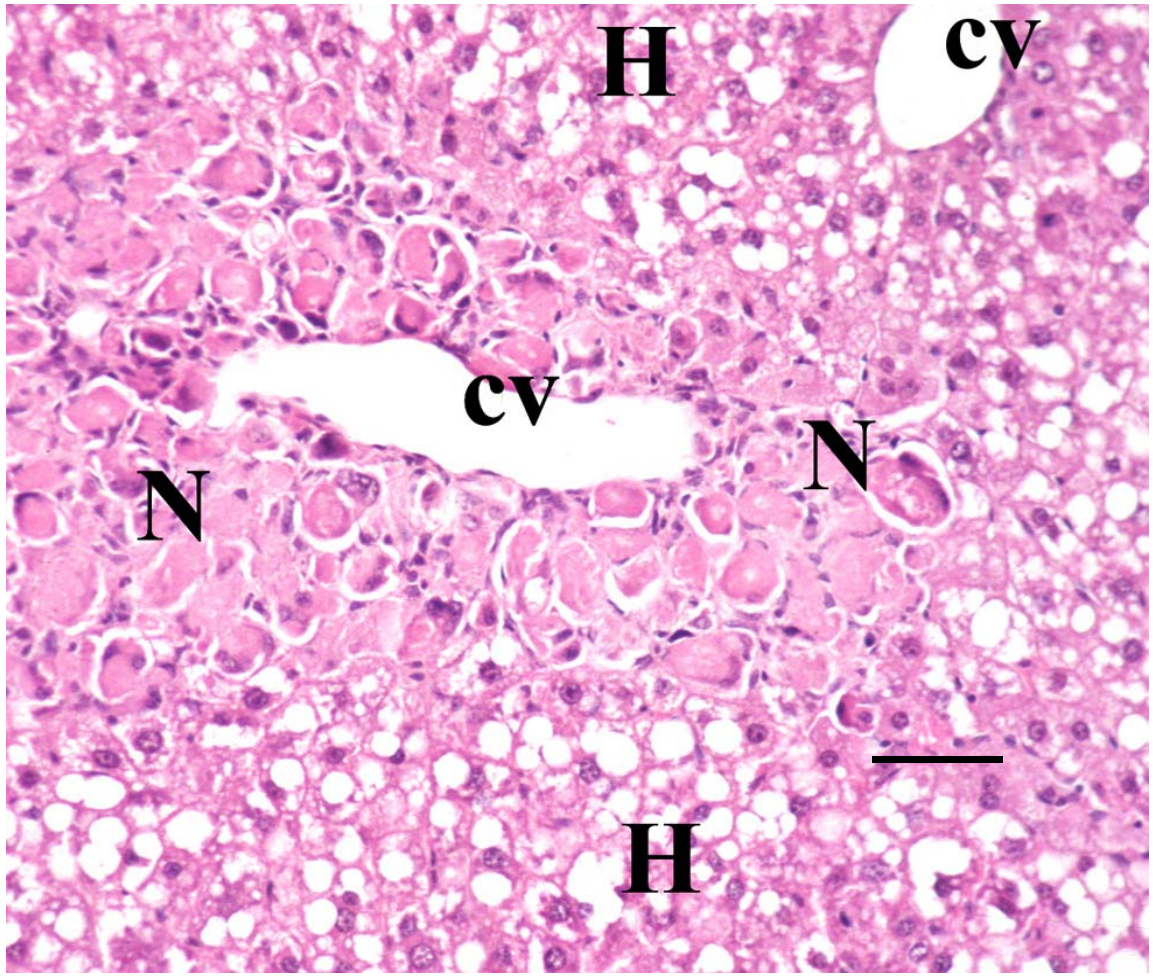


Figure 4. Example of severity category 3. Liver of a mouse treated with BDCM (3.05 mmol/Kg) showing severe centrilobular necrosis (N), hepatocytes (H), and central vein (CV). H&E, bar = 40 μm.

Table 3
Model 1 Results: Prediction of Pathology Scores from Dose

Logistic Model: $\ln(\pi_i/(1-\pi_i)) = \alpha_i + \beta_1$ Dose

Significant Test for the Proportional Odds Assumption:
Chi-Square = 0.0063 with 1 DF (p=0.9369)
(Null hypothesis of equal slopes is not rejected.)

Significant Tests of Overall Model (Null Hypothesis $\beta_1=0$):
-2 LOG L = 90.016 with 1 DF (p=0.0001)
Score = 45.312 with 1 DF (p=0.0001)
(Null hypothesis that overall model does not explain the severity category is rejected.)

Significant Chi-Square Tests of Null Hypotheses for Parameter Estimates:

Variable	Estimate	Null Ho:	DF	P-value
Intercept 1	5.5	$\alpha_1=0$	1	0.0008
Intercept 2	14.6	$\alpha_2=0$	1	0.0003
Dose	-9.33	$\beta_1=0$	1	0.0003

(Null hypotheses that the intercept terms and dose individually do not explain the severity category are rejected.)

Table 4
Model 2 Results: Prediction of Pathology Scores from ln(ALT) and PcLiv

Logistic Model:

$\ln(\pi_i/(1-\pi_i)) = \alpha_i + \beta_1 \ln(ALT) + \beta_2(PcLiv)$

Significant Test for the Proportional Odds Assumption:

(Null hypothesis of equal slopes is not rejected.)
Chi-Square = 0.3888 with 2 DF (p=0.8233)

Significant Tests of Overall Model (Null Hypothesis $\beta_1=\beta_2=0$):

-2 LOG L = 65.548 with 2 DF (p=0.0001)
Score = 43.203 with 2 DF (p=0.0001)
(Null hypothesis that overall model does not explain the severity category is rejected.)

Significant Chi-Square Tests of Null Hypotheses for Parameter Estimates:

Variable	Estimate	Null Ho:	DF	P-value
Intercept 1	12.5	$\alpha_1=0$	1	0.0001
Intercept 2	16.9	$\alpha_2=0$	1	0.0001
PCLiv	-1.2	$\beta_1=0$	1	0.0095
Ln(ALT)	-3.7	$\beta_2=0$	1	0.0056

(Null hypotheses that the intercept terms, PcLin and Ln(ALT), individually do not explain the severity category are rejected.)

Table 5
Model 1 Results: Probability of Observing a Severity Level Given Dose

Dose (mg/kg)	Dose (mmol/kg)	**Low95**	**P(S>1)** (> None)	**UPP95**	P(S=1) (None)	P(S=2) (Mild-Mod)	P(S=3) (Mod-Sev)
0.0	0.00	0.00016	0.00396	0.09142	0.99604	0.00396	0.00000
10.4	0.05	0.00032	0.00630	0.11280	0.99370	0.00630	0.00000
20.8*	**0.10**	**0.00063**	**0.01001**	**0.13888**	**0.98999**	**0.01000**	**0.00000**
31.2	0.15	0.00126	0.01586	0.17052	0.98414	0.01586	0.00000
41.7	0.20	0.00250	0.02505	0.20865	0.97495	0.02505	0.00000
52.0	0.25	0.00490	0.03936	0.25425	0.96064	0.03936	0.00000
62.5	0.30	0.00949	0.06133	0.30827	0.93867	0.06132	0.00001
72.9	0.35	0.01803	0.09436	0.37148	0.90564	0.09434	0.00001
83.3	0.40	0.03337	0.14246	0.44427	0.85754	0.14244	0.00002
93.7	0.45	0.05945	0.20943	0.52612	0.79057	0.20940	0.00003
104.0	0.50	0.10053	0.29696	0.61485	0.70304	0.29692	0.00005
114.6	0.55	0.15916	0.40246	0.70559	0.59754	0.40239	0.00008
125.0	0.60	0.23378	0.51784	0.79082	0.48216	0.51771	0.00012
135.4	0.65	0.31857	0.63134	0.86251	0.36866	0.63114	0.00020
145.8	0.70	0.40614	0.73196	0.91599	0.26804	0.73164	0.00032
156.2	0.75	0.49053	0.81323	0.95167	0.18677	0.81273	0.00050
166.6	0.80	0.56816	0.87410	0.97343	0.12590	0.87330	0.00080
177.0	0.85	0.63745	0.91716	0.98586	0.08284	0.91588	0.00128
187.5	0.90	0.69799	0.94639	0.99264	0.05361	0.94435	0.00204
197.9	0.95	0.75006	0.96569	0.99623	0.03431	0.96245	0.00324
208.3	1.00	0.79429	0.97821	0.99809	0.02179	0.97304	0.00516
218.7	1.05	0.83147	0.98622	0.99904	0.01378	0.97801	0.00821
229.1	1.10	0.86245	0.99131	0.99952	0.00869	0.97829	0.01303
239.5	1.15	0.88809	0.99453	0.99976	0.00547	0.97392	0.02061
250.0	1.20	0.90917	0.99657	0.99988	0.00343	0.96410	0.03247
260.4	1.25	0.92643	0.99784	0.99994	0.00216	0.94705	0.05079
270.8	1.30	0.94049	0.99865	0.99997	0.00135	0.92003	0.07862
281.2	1.35	0.95193	0.99915	0.99999	0.00085	0.87939	0.11976
291.6	1.40	0.96120	0.99947	0.99999	0.00053	0.82120	0.17827

* Dose approximately equal to a LOAEL of 17.9 mg/kg/day for renal cytomegaly in $B6C3F_1$ male mice (NTP, 1986)

shown is the probability of observing any adverse or severe effect [P(S=2 or 3) or equivalently, P(S>1)]. This translates into the probability of observing any centrilobular necrosis at all, given dose. The RfD for BDCM (U.S. EPA, 1999) was calculated from an animal chronic LOAEL of 17.9 mg/kg/day (NTP, 1986) based on renal cytomegaly in a corn oil, gavage bioassay on B6C3F1 male mice. At approximately this LOAEL of 17.9 mg/kg (20.8 mg/kg in Table 5), *model 1* estimates that the probability of observing any centrilobular necrosis in this data set [P(S>1)] is 0.01, with a 95% confidence interval of 6E-4 to 0.14. This is consistent with the fact that the kidney effect chosen for the RfD development should be a more sensitive endpoint than the liver endpoint.

Model 2

Table 6 shows some predicted probabilities generated from *model 2*, which was developed by regressing severity categories on two explanatory variables, ln(ALT) and PcLiv, independent of dose. The eight point data set shown here (all mice dosed with 1.2 mmol/kg/d of BDCM) is a test data set not used to develop *model 2*, but rather is presented to illustrate the usefulness of such a model. The larger probabilities (in bold) can be used as indicators of the occurrence and severity of centrilobular necrosis, given information only on ALT and percent relative liver weight. For occurrence, seven of the eight observed pathology severity scores (of either 2 or 3) match the prediction of at least mild centrilobular necrosis (i.e., larger probabilities are for severities 2 or 3). For severity, in four of the eight mice, the observed pathology severity score corresponded exactly to the severity of the necrosis associated with the largest probability. In one case, neither occurrence nor severity was correct, as the predicted severities and observed pathology results did not agree.

Conclusions

The CR procedures demonstrated here with *models 1 and 2* are a useful way to evaluate categorical and qualitative histopathology data. Models, such as *model 1*, offer a way to compare new toxicity results with existing benchmarks, such as RfDs, RfCs, no-effect levels, benchmark dose modeling results, incidence levels for certain endpoints of interest, or doses that produce specific levels of severity seen in other studies or other species. Such comparisons allow the investigator to judge new findings against other existing data to aid in the interpretation of experimental results. The results of this investigation compare the LOAEL used to calculate the BDCM RfD with the hepatotoxic data in the 14-day gavage study in female CD-1 mice. This comparison supports the use of nephrotoxic data from the chronic bioassay in male B6C3F1 mice as the appropriate critical effect for the calculation of the RfD for BDCM because the CR analysis seems to indicate the liver data are less sensitive.

The predictions of occurrence and severity of centrilobular necrosis, such as those made using *model 2*, from liver enzyme and weight data are promising for use in the initial screening of hepatotoxic data. In this investigation, the model predicted the occurrence of centrilobular

Table 6

Model 2 Results: Predicted Probability of Centrilobular Necrosis
for Bromodichloromethane at1.2 mmol/k/d Exposure

Mouse ID #	ln(ALT) IU/l	% Relative Liver Weight	Predicted P(S=1) None Detected	Predicted P(S=2) Mild to Mod	Predicted P(S=3) Mod-Sev to Sev	Observed Pathology Severity Score*	Pathology Results** (Examined Blind)
13	1.97	6.34	0.07	**0.79**	0.14	3	Mod-sev cenlob vacuolation Mod-sev cenlob necrosis Mod zone 2 vacuolation
19	2.75	6.07	0.006	0.31	**0.68**	2	Mod cenlob necrosis Mod cenlob vacuolation
36	2.23	6.89	0.01	**0.52**	**0.47**	2	Mod cenlob necrosis Mild-mod cenlob vacuolation
50	2.30	6.24	0.02	**0.64**	0.34	1	Mod cenlob vacuolation Mod zone 2 vacuolation
53	2.23	6.80	0.01	**0.54**	**0.44**	2	Mod cenlob necrosis Mod cenlob vacuolation Mild-mod zone 2 vacuolation

Table 6 cont.							
Mouse ID #	ln(ALT) IU/l	% Relative Liver Weight	Predicted P(S=1) None Detected	Predicted P(S=2) Mild to Mod	Predicted P(S=3) Mod-Sev to Sev	Observed Pathology Severity Score*	Pathology Results** (Examined Blind)
69	2.29	6.82	0.01	**0.48**	**0.51**	3	Sev cenlob necrosis Sev cenlob vacuolation Mod zone 2 vacuolation Few mitosis
85	2.55	5.62	0.02	**0.61**	0.37	3	Mod-sev cenlob necrosis Mod-sev cenlob vacuolation Mod zone 2 vacuolation Granuloma Mod mitosis
90	1.79	6.01	0.17	**0.77**	0.05	2	Mod cenlob necrosis Mod mitosis Mild cenlob necrosis

*Pathology score for centrilobular necrosis: 1=none; 2=mild, mild-moderate, moderate; 3= moderate-severe, severe
**Abbreviations: Mild = mild; Mod=Moderate; Sev=Severe; Cenlob=Centrilobular

necrosis well for the test data set, but was less accurate in predicting the severity of the effect. Additional research is needed to expand the development and testing of this model using other data sets. Refinement and standardization of such models in the future could also reduce the need to remove and weigh livers from every animal in a study or to perform pathology examinations on all liver tissues. If predictions of liver pathology could be made reliably, then a random sample of liver tissues could be evaluated, with the results checked against the CR model for consistency. A related modeling effort would be to predict centrilobular necrosis using enzyme data alone. Future efforts will concentrate on including additional enzyme levels or other variables in the model and on investigating the use of such models across similar chemicals within a class of compounds.

References

Dourson, M.L., L.K. Teuschler, P.R. Durkin and W.M. Stiteler. 1997. Categorical regression of toxicity data: A case study using aldicarb. *Reg. Tox. And Pharm:* 25:121-129.

Farland, W. and M.L. Dourson. 1992. Noncancer health endpoints: Approaches to quantitative risk assessment. In: Comparative Environmental Risk Assessment. R. Cothern, (ed.) Boca Raton, FL: Lewis Publishers Inc.

Guth, D.J., A.M. Jarabek, L. Wymer and R.C. Hertzberg. 1991. Evaluation of risk assessment methods for short-term inhalation exposure. Presented at the 84th Annual Meeting of the Air & Waste Management Association, Vancouver, British Columbia, June 16-21.

Guth, D.J., R.J. Carroll, D.G. Simpson, and H. Zhou. 1997. Categorical regression analysis of acute exposure to tetrachloroethylene. *Risk Analysis*. 17(3):321-332.

Hertzberg, R.C. 1989. Extrapolation and scaling of animal data to humans: Fitting a model to categorical response data with application to species extrapolation of toxicity. *Health Physics.* 57(Sup 1), 405-409.

Hertzberg, R.C. and M. Miller. 1985. A statistical model for species extrapolating using categorical response data. *Toxicol. Indus. Health.* 1(4), 43-63.

NTP (National Toxicology Program). 1986. Toxicology and carcinogenesis studies of bromodichloromethane in F344/N Rats and B6C3F1 Mice (gavage studies). NTP Technical Report, Ser. No. 321, NIH Publ. No. 87-2537.

Plaa, G.L. and W.R. Hewitt. 1994. Detection and evaluation of chemically induced liver injury. In: Principles and Methods of Toxicology, Second Edition. A. Wallace Hayes (ed.). New York: Raven Press, Ltd., 23:839-870.

Rao, V.R., K. Levy and M. Lustik. 1993. Logistic regression of inhalation toxicities of perchloroethylene - Application in noncancer risk assessment. *Reg. Toxicol. Pharmacol.* 18, 233-247.

SAS Logistic Procedure. 1990. SAS Institute, Inc. Cary, NC.

U.S. EPA. 1999. Integrated Risk Information System (IRIS). Online. National Center for Environmental Assessment, Washington, DC.

A PROPORTIONAL-RESPONSE ADDITION MODEL FOR EVALUATING HEALTH RISKS ASSOCIATED WITH COMPLEX MIXTURES: AN EXAMPLE USING DRINKING WATER DISINFECTION BY-PRODUCTS

William M. Stiteler[1], Linda K. Teuschler[2], Joan T. Colman[1], Richard C. Hertzberg[2]

[1]Syracuse Research Corporation
[2]U.S. EPA, National Center for Environmental Assessment - Cincinnati Office

Abbreviated Title: Additivity Model for Complex Mixtures

September 2000

Abstract

Complex mixtures such as those resulting from disinfection of drinking water supplies present a difficult problem for risk assessors because of the large number of components. The concentration of each individual component may be small enough to pose no threat, but if they act jointly in some additive or greater than additive manner, then there may be some concern for health safety.

For the most part, models for the joint action of chemical mixtures are based on assumptions about the mechanisms of action. For example, response addition is based on the assumption of independent but joint action, while dose addition is based on an assumption of simple similar action. This reliance on mechanism becomes problematic for complex mixtures. As the number of components becomes large, it is unlikely that a single mechanism-based definition of additivity would be universally applicable across the whole mixture. There are, for example, 1,225 different pairs of components in a 50-component mixture. Different pairs of components may exhibit different modes of action. Even if a common mode of action can be assumed across the mixture, the usual definitions of additivity may not be appropriate. If all components are present at sub-threshold levels, then response addition will indicate no response for the mixture, regardless of how many components there are.

The solution proposed in this paper is to use proportional-response addition, a generic definition of additivity that does not depend on mechanism of action, for modeling complex mixtures. An example of the application of this model to a mixture of DBPs is presented.

Introduction

The risk assessment community has evaluated and addressed issues related to remediation for many of the obvious, major pollutants in the environment (e.g., from tobacco smoke, dioxin, or mercury exposure). The health hazards from other pollutants are often less obvious because of many factors, such as variations in occurrence, inadequate analytical tools for detection and characterization, or subtle and uncertain health effects at low exposures to multiple stressors. Because the risk assessment frontier is approaching scenarios that involve multi-route and multi-chemical exposures that may vary over time or with environmental conditions, new risk assessment methodologies are needed. For chemical mixtures risk assessment, a number of new approaches are being developed, each of which has its own objectives and assumptions [1,2]. To implement many of these methodologies, a "best guess" must be made relative to the toxicologic mechanism of action for the combination of chemicals, and based on this judgment, specific statistical methods for combining dose-response data with exposure estimates can be used to present a risk characterization.[3]

Existing component-based methods that assess the risk of mixtures make assumptions about the joint action of the chemicals (i.e., usually dose additivity or response additivity). Response addition is based on the assumption of independent but joint action, while dose addition is based on an assumption of simple similar action.[4] While an assumption of "additivity" may be reasonable in the low-dose region (where there is little or no toxicologic response), it is unlikely that a single mechanism-based definition of additivity would be universally applicable across the whole mixture. Different pairs of components may exhibit different modes of action. For example, an evaluation of mechanisms of action might suggest that dose additivity is appropriate for components A and B and that response additivity is appropriate for components B and C. For some pairs of components, the mechanism of action may be unknown. The approach that one would use to incorporate information about individual components into an estimate of risk for a binary mixture would depend on the mechanism associated with the type of additivity that is assumed. As a result, for a complex mixture, it may not be prudent to base the methodology for estimating risk on a definition of additivity that requires a specific mechanism of action. This paper introduces the use of another definition of additivity, proportional-response additivity, that does not rely on an assumption about mechanism; this new approach is illustrated using the complex mixture of disinfection by-products (DBPs) in drinking water.

Public drinking water supplies are generally treated to reduce health threats from waterborne pathogens. These treatments involve a variety of processes including filtration, aeration, and disinfection. The disinfectants operate through reactive processes that produce, and leave in the water, a large number of chlorinated or halogenated DBPs. To date, several hundred DBPs have been identified, the most common of which belong to a few distinct chemical classes: trihalomethanes, haloacetonitriles, haloacetic acids, chlorophenols, chlorinated ketones, chlorinated furanones, and chlorinated aldehydes.[5,6,7,8,9]

Consumption of the finished drinking water exposes the human population to a complex mixture of these DBPs. The exact composition of the mixture varies depending on the selection and sequencing of the various processes that make up the treatment process. To further complicate the problem, the composition of the DBP mixture also varies, both temporally and spatially, for a particular treatment process, depending on a number of factors, including source water characteristics (e.g., temperature, pH, etc.) and distribution system effects (e.g., regrowth of microbes, chlorine injection, distance from the treatment plant, etc.).

Individual DBPs are generally present in drinking water at levels low enough that, by themselves, they may pose no serious threat to human health. This is particularly true for endpoints such as reproductive or developmental effects, for which there is consideration of a toxicity threshold. However, some of the DBPs are associated with adverse health effects at higher concentrations in animal studies [10,11,12,13,14]. This, coupled with the fact that humans are exposed to DBPs as a complex mixture, is cause for concern about their combined effect. If DBPs act jointly in some additive or greater than additive manner, then there may be some concern for health safety even when individual concentrations may not be cause for concern. The challenge facing the risk assessment community today is to characterize both the magnitude and the potential severity of any health risk than can be associated with these complex mixtures of DBPs.

Methods

It is proposed that a generic definition of additivity, independent of mechanisms of action, be used as the default. One such definition of additivity, referred to here as "proportional- response addition", has been widely applied in many fields including industrial engineering and agriculture [15,16]. It is a definition that arises intuitively as the following example illustrates. Suppose that a particular automobile gets 30 miles per gallon of Brand A gasoline and only 20 miles per gallon of Brand B gasoline. What would be the "expected" mileage if a mixture of equal parts of A and B is used in that automobile (assuming other factors are held constant)? The answer most people would give to that question would be 25 miles per gallon. Anything above 25 would be considered greater than additive and anything smaller than 25 would be considered less than additive. In spite of this intuitive basis for proportional-response addition, it has not been widely applied in toxicology. This is probably because much of the initial mixtures work has focused on binary mixtures where mechanism of action would be a natural consideration in defining and classifying joint action.

One exception is a paper by Chen et al.[17] that describes proportional-response addition within a toxicological context and, in fact, suggests that proportional-response addition should be used as a substitute for the "usual" definition of response addition. Others have applied the definition[18,19] where it appears to have been based on an intuitive notion of what additive means rather than on any deliberate selection from among various definitions.

Figure 1 illustrates the concept behind proportional-response addition for a binary mixture. If the response to some specified dose is known for both chemical A and chemical B,

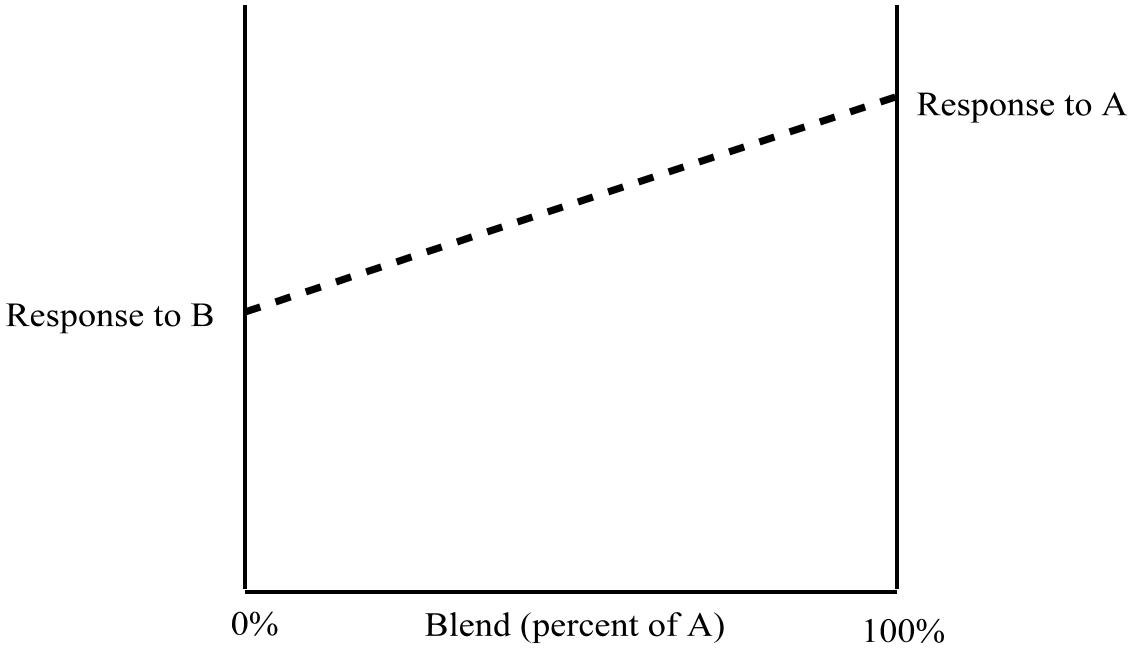


Figure 1. Proportional-Response Addition.

then the "expected" response to various mixture blends of A and B, keeping the total dose fixed, might be estimated from the relative proportions of the components in that mixture. The dotted line in Figure 1 illustrates how the expected response would vary as the proportion of the total fixed dose varies in composition from 0% of A to 100% of A.

This definition of additivity is easily generalized for any number of mixture components. Suppose a mixture is composed of n components and the dose-response relationship for the i^{th} component is represented by P_i(dose). Let the individual doses of the n components be represented by d_1, d_2, d_3, ...,d_n. The total amount or dose of the mixture is then $D = d_1 + d_2 + ... + d_n$. The proportion of that total amount represented by component i is then given by $\pi_i = d_i / D$. Then, under proportional-response addition, the response to the total dose, D, of the mixture would be:

$$P_{mix}(D) = \pi_1 \cdot P_1(D) + \pi_2 \cdot P_2(D) + \cdots + \pi_n \cdot P_n(D). \quad (1)$$

Figure 2 shows the dose-response surface for a hypothetical binary mixture at varying mixture total doses and mixture blends. As another example, suppose there are four components in a mixture and the amounts are 10, 15, 25, and 50 mg/kg, respectively. Then the total amount of the mixture would be D = 100 mg/kg and the relative proportions of the four components would be 0.1, 0.15, 0.25, and 0.5, respectively. Assuming individual dose response curves P_1, P_2, P_3, and P_4, are available, the response to the mixture under proportional-response addition would be calculated as:

$$P_{Mix}(100) = 0.1 \cdot P_1(100) + 0.15 \cdot P_2(100) + 0.25 \cdot P_3(100) + 0.5 \cdot P_4(100). \quad (2)$$

This approach involves estimating the risk for each individual component at the level of the whole mixture. This gives an estimate of the risk for each component as if it were present at a level equal to the whole mixture. These risks are then apportioned according to the percentage of the total mixture that each component represents. Thus, extrapolation into the low dose region is avoided; this is especially important for those effects that may have thresholds.

Relationship between proportional-response addition and response addition

It is of interest to explore the relationship between proportional-response addition and the usual definition of response addition. Response addition assumes the components of the mixture act independently of one another, both toxicologically and under the assumption of statistical independence. Under certain conditions, the two definitions give equivalent results. This occurs when: 1) the dose response relationship is linear and there are no thresholds; and 2) there is no background response or the risk is calculated as extra risk so the intercept of the dose response line is zero. Under these conditions, the contribution of chemical 1 to the total risk under proportional-response addition would be $[d_1 /(d_1 + d_2)]\cdot P_1(d_1 + d_2)$, while under response addition,

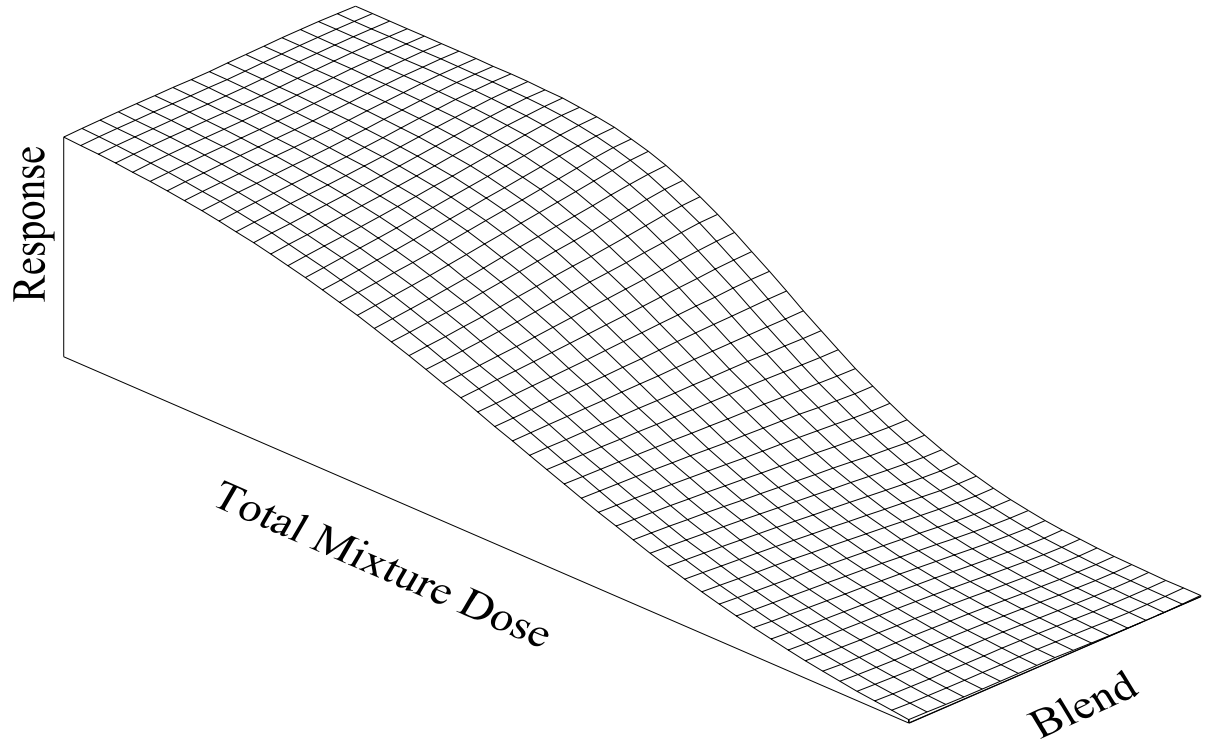


Figure 2. Dose-Response Surface for Binary Mixture.

the contribution of chemical 1 to the total risk would be $P_1(d_1)$. These are equal, however, if the dose-response relationship is linear and has no threshold as Figure 3 illustrates. The slope of the line in Figure 3 is given by:

$$\text{Slope} = \frac{P_1(d_1)}{d_1} = \frac{P_1(d_1 + d_2)}{d_1 + d_2}. \quad (3)$$

Then, multiplying by d_1 gives the result. The contribution of chemical 2 would be calculated in a similar manner.

Note that if the dose-response function for the first component has a threshold, then it is possible that $P_1(d_1) = 0.0$ while $P_1(d_1 + d_2) > 0.0$. In that case, the two methods lead to different estimates of risk.

Thus, under conditions that might be encountered with a non-threshold endpoint (e.g., cancer) where exposures are in the linear portion of the dose-response curve, proportional-response addition and response addition would result in identical risk estimates.

Results

The developmental and reproductive health risks for two classes of DBPs, the haloacetic acids and the haloacetonitriles, were evaluated as an illustration of the proportional-response additivity methodology. The haloacetic acids and haloacetonitriles commonly found in drinking water are listed in Table 1 (see References in Table 2). Table 1 shows the chemical name, formula, and acronyms used in this presentation. In addition, the table presents an indication of the availability of developmental and reproductive toxicity data. Seven of the eleven haloacetic acids (MCA, DCA, TCA, MCA) and haloacetonitriles (DCAN, TCAN, BCAN) have been subjects of developmental toxicity studies by a single group of investigators, and three of the haloacetic acids (DCA, MBA, DBA) have been the subjects of male reproductive studies by another group of investigators. These studies all were conducted in rats using gavage administration. The results for developmental toxicity were positive. For reproductive toxicity, the dihalogenated haloacetic acids gave positive results, but monohalogenated acetic acid (MBA) gave negative results. An additional chemical, the haloacetonitrile DBAN, was tested in a short-term developmental and reproductive toxicity screening study in rats, with negative results. DBAN was administered in the drinking water, which is preferable in terms of relevance to human exposure to DBPs. The animals, however, refused to drink the DBAN solutions at higher concentrations, so the maximum tolerated dose was defined on the basis of taste aversion rather than actual toxicity. The negative results from that study are present in parentheses, to indicate this was a screening-level study and a toxicity-based MTD was not achieved. In addition, the positive results for MCA were borderline in terms of statistical significance and are therefore shown in parentheses.

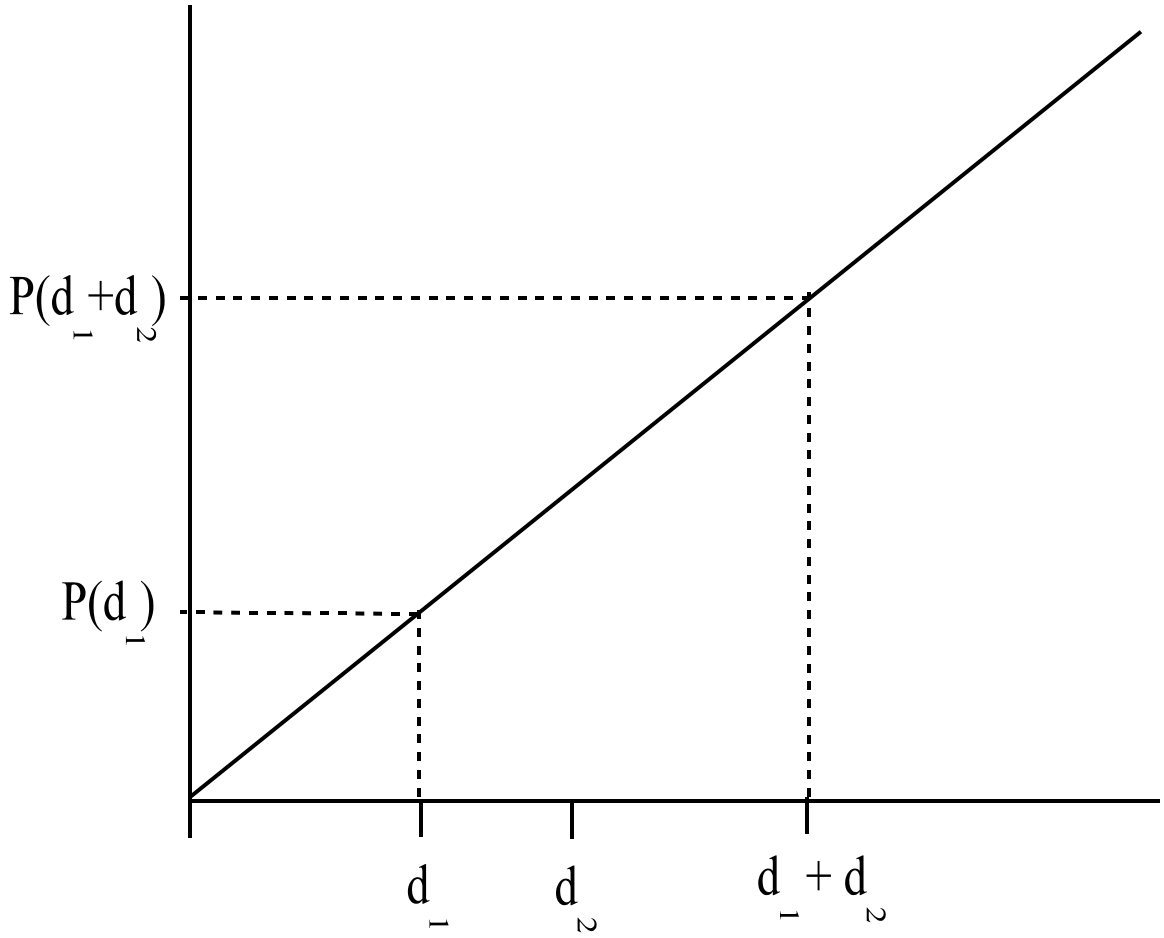


Figure 3. Linear model evaluated at d_1 and at $d_1 + d_2$.

Table 1. Availability of Developmental and Reproductive Dose-Response Data for Haloacetic Acids and Haloacetonitriles Found in Drinking Water.

Chemical			Developmental Toxicity[a]	Reproductive Toxicity[a]
Haloacetic Acids				
$ClCH_2COOH$	Monochloroacetic Acid	MCA	y, (+)	
$Cl_2CHCOOH$	Dichloroacetic Acid	DCA	y, +	y, +
Cl_3CCOOH	Trichloroacetic Acid	TCA	y, +	
$BrCH_2COOH$	Monobromoacetic Acid	MBA	y, +	y, -
$Br_2CHCOOH$	Dibromoacetic Acid	DBA		y, +
BrClCHCOOH	Bromochloroacetic Acid	BCA		
Haloacetonitriles				
Cl_2CHCN	Dichloroacetonitrile	DCAN	y, +	
Cl_3CCN	Trichloroacetonitrile	TCAN	y, +	
BrClCHCN	Bromochloroacetonitrile	BCAN	y, +	
Br_2CHCN	Dibromoacetonitrile	DBAN	y,(-)[b]	y, (-)[b]
Br_2ClCCN	Dibromochloroacetonitrile	DBCAN		

[a]Data are from gavage studies in rats unless otherwise noted.
[b]Data are from a screening-level drinking water study in rats.
y = yes, adequate data available
+ = results were positive for adverse effect
- = results were negative for adverse effect
(+) = results were marginally positive
(-) = results were negative, but a toxicity-based MTD could not be achieved because of taste aversion and consequent refusal to drink higher concentrations of the chemical, and this was a short-term screening study

The use of common sensitive endpoints across chemicals was considered desirable because it supported a greater variety of modeling approaches, including the possibility of combining data. Data sets were examined with regard to NOAELs and LOAELs to select endpoints for each compound that appeared sensitive and were reported across the entire series of compounds. Table 2 lists the endpoints further examined using a threshold dose-response model as described subsequently in this section. For the purposes of this illustration, focus was on the developmental toxicity data because more chemicals were tested, and all the chemicals tested by gavage, including both chlorinated and brominated compounds and both haloacetic acids and haloacetonitrile, gave positive results. Thus, it appeared that developmental toxicity was characteristic of these chemical classes.

Further evaluation of the data sets listed in Table 2 by dose-response modeling (discussed in next paragraph) showed that visceral malformations, particularly cardiovascular (interventricular septal defects, defect between ascending aorta and right ventricle, and levocardia) and smaller fetal size (body weight and crown-rump length) appeared to be the most sensitive endpoints in common for these chemicals. An example of these data are shown for DCA in Table 3. Note that some of the data are quantal, but other data (body weight, crown-rump length) are continuous, and were converted to quantal (*estimated # of litters affected* in the table) prior to modeling.

Conversion of the continuous-response developmental data to quantal form was performed by assuming a normal distribution with a constant variance across dose groups for the response and 5% background response rate. Because individual animal data were not available, the number of responders in each dose group was estimated by first establishing a critical value representing the point above (or below, depending on the direction of adverse response) which 5% of the control group lies. Then, for each dose group, the proportion exceeding this critical value was estimated. This proportion was applied to the number of animals in the dose group to determine the number of responders. The doses were converted to equivalent human doses using the scaling factor of body weight to the 3/4 power. A threshold dose-response model was then fit to the data sets for the individual DBPs.

As shown previously in Table 2, adequate developmental toxicity data were lacking for the haloacetic acids DBA and BCA and for the haloacetonitriles DBAN and DBCAN. A surrogate approach seemed appropriate to fill these data gaps, because the available data indicated that developmental toxicity may be common to the haloacetic acid and haloacetonitrile DBPs. As a provisional measure, DCA was selected as a surrogate for these haloacetic acids and TCAN was selected as a surrogate for these haloacetonitriles. A search for mechanistic data to support selection of surrogates revealed studies of mechanisms relevant to carcinogenicity, which may not be relevant to developmental toxicity, and some *in vitro* embryo culture studies, which did not appear to give results corresponding to the available *in vivo* testing. Therefore, the selections of surrogates were based partly on structural similarity and partly on quality of the data (such as better dose spacing) for the surrogate.

Table 2. Developmental and Reproductive Toxicity Data Sets for Haloacetic Acids and Haloacetonitriles.

Chemical	Study	Endpoint (in rats except as noted)
Haloacetic Acids - Developmental Toxicity		
Monochloroacetic Acid	Smith et al.[20]	Fetal body weight
		Crown-rump length
		Visceral malformations, total (% affected/litter
Dichloroacetic Acid	Smith et al.[21]	Fetal body weight - male
		Fetal body weight - female
		Crown-rump length - male
		Crown-rump length - female
		Visceral malformations, total
		Visceral malformations, cardiovascular
Trichloroacetic Acid	Smith et al.[14]	Complete litter resorption
		% Postimplantation loss/litter
		Fetal body weight - male
		Fetal body weight - female
		Fetal crown-rump length - male
		Fetal crown-rump length - female
		Visceral malformations, total
		Visceral malformations, cardiovascular
		Visceral malformations, levocardia
		Skeletal malformations

Chemical	Study	Endpoint (in rats except as noted)
Monobromoacetic Acid	Randall et al.[22]	Fetal body weight
		Fetal crown-rump length
		Visceral malformations, total (% affected/litter)
Haloacetic Acids - Male Reproductive Toxicity		
Dichloroacetic Acid	Cicmanec et al.[23]	Testicular lesions: degeneration
	Linder et al.[24]	Number caput sperm
		Number cauda sperm
		% Motile sperm
		Progressive motility
		Testicular histopathology: Faulty spermiation
Dibromoacetic Acid	Linder et al.[25]	Number caput sperm
		Number cauda sperm
		% Motile sperm
		Progressive motility
		Retention Stage IX spermatids per tubule

Chemical	Study	Endpoint (in rats except as noted)
Haloacetonitriles - Developmental Toxicity		
Dichloroacetonitrile	Smith et al.[26]	Complete litter resorption
		% Postimplantation loss/litter
		Fetal body weight - male
		Fetal body weight - female
		Fetal crown-rump length - male
		Fetal crown-rump length - female
		Visceral malformations, total
		Visceral malformations, cardiovascular
		Visceral malformations, urogenital
		Skeletal malformations
Trichloroacetonitrile	Smith et al.[27]	Complete litter resorption
		% Postimplantation loss/litter
		Fetal body weight - male
		Fetal body weight - female
		Visceral malformations, total
		Visceral malformations, cardiovascular
		Visceral malformations, urogenital

Chemical	Study	Endpoint (in rats except as noted)
Bromochloroacetonitrile	Christ et al.[28]	Complete litter resorption
		% Postimplantation loss/litter
		Fetal body weight - male
		Fetal body weight - female
		Fetal crown-rump length - male
		Fetal crown-rump length - female
		Visceral malformations, total
		Visceral malformations, cardiovascular
		Visceral malformations, urogenital
		Skeletal malformations

Table 3. Example Data Sets: DCA (Smith et al.[21]).

Dose in mg/kg-day on gestation days 6-15			0	14	140	400
Fetal body weight (g) (mean of litter means)						
# litters examined			19	18	19	19
male	mean		3.68	3.75	3.6	3.43
	SD		0.2	0.3	0.2	0.3
	statistical significance					*
	estimated # litters affected[a]		1	0	2	5
female	mean		3.49	3.6	3.46	3.27
	SD		0.2	0.3	0.2	0.3
	statistical significance					*
	estimated # litters affected[a]		1	0	1	4
Fetal crown-rump length (cm) (mean of litter means)						
# litters examined			19	18	19	19
male	mean		3.62	3.64	3.56	3.46
	SD		0.1	0.2	0.1	0.2
	statistical significance					*
	estimated # litters affected[a]		1	1	2	5
female	mean		3.55	3.59	3.49	3.38
	SD		0.1	0.2	0.1	0.2
	statistical significance					*
	estimated # litters affected[a]		1	1	2	5
Visceral malformations:						
# litters examined[b]			39	18	19	19
Total visceral						
# litters affected			0	1	4	7
% fetuses affected/litter						
	mean		0	0.69	2.6	9.82
	SD			2.95	5.6	17.2
	statistical significance				*	*
Cardiovascular						
# litters affected			0	1	2	6
% fetuses affected/litter						
	mean		0	0.69	1.02	8.07
	SD			2.95	3.1	16.26
	statistical significance					*

[a]Continuous data were converted to quantal form as described in the text
[b]For controls, # of litters examined is from two related studies, combined
*$p \leq 0.05$

The proportional-response additivity approach was applied to concentration data from two pilot studies of drinking water disinfection processes: one using Mississippi River water at Jefferson Parish, LA (Lykins et al.[7]) and the other using Ohio River water (Miltner et al.[8]; with further analysis by NCEA-CIN EPA). Data sets for water treated with chlorine were chosen because this process results in higher levels of the haloacetic acids and haloacetonitriles relative to other methods, such as pretreatment with ozone prior to disinfection with chlorine, and/or other disinfectants such as chloramine.

Discussion and Conclusions

Risks of developmental toxicity were estimated for humans ingesting drinking water containing the haloacetic acids and haloacetonitriles at the concentrations determined in the pilot studies described in the previous section. The risk estimates were based on increased total visceral malformations, increased cardiovascular malformations, decreased fetal body weight and decreased fetal crown-rump length. Because the reports of the animal data for MCA and MBA did not break out the cardiovascular incidence data from the total visceral malformation data (stated to be mainly cardiovascular), the proportional risk estimates for cardiovascular malformations used total visceral malformations for those two chemicals. The risk estimates were similar for total visceral malformations, cardiovascular malformations, and decreased fetal weight, and somewhat lower for decreased fetal crown-rump length.

The risk estimates based on total visceral malformations were selected to illustrate the application of proportional-response additivity, because these estimates best illustrated the use of surrogates. Details of these estimates are provided in Tables 4 and 5. The concentration data, in μg/L, were converted to human doses by assuming water consumption of 2 L/day and a body weight of 70 kg. The total combined dose of haloacetic acids and haloacetonitriles was approximately 4 μg/kg-day for Jefferson Parish (Mississippi River) water and 2 μg/kg-day for the Ohio River water. The risk for each DBP was estimated at the total mixture dose (as if the DBP were present at a dose equal to the whole mixture dose) using a threshold dose response model as previously described. The risk for some components was zero because the thresholds for these components were higher than the total mixture dose. The proportional risk for each component was then calculated by multiplying risk at the total mixture dose by the proportion of that component in the mixture. The sum of the proportional risks, or total risk, for the two sets of drinking water data was 1.2×10^{-5} for Jefferson Parish/Mississippi River and 8.9×10^{-6} for the Ohio River. These values are virtually identical, and indicate a relatively low risk for developmental toxicity (≈1 in 100,000) during the gestational period.

If proportional risks based on surrogates were omitted, the total risks for the two sets of drinking water data were slightly lower: 7.4×10^{-6} and 7.1×10^{-6}, respectively (0.7 in 100,000). Elimination of surrogates from the risk estimates had a similar impact for estimates based on cardiovascular malformations, but virtually no impact on the estimates based on fetal body weight and none on crown-rump length. Risk was also estimated for an alternative disinfection method involving pretreatment of Ohio River water with ozone followed by chlorine. This

Table 4. Example Risk Estimate for Developmental Toxicity in Humans Exposed to Haloacetic Acids and Haloacetonitriles in Drinking Water at Concentrations Determined in Pilot Study, Jefferson Parish (Mississippi River) Following Chlorine Treatment.

Chemical	Water Concentration µg/L	Estimated Dose mg/kg-day	Proportion of Component in Mixture	Risk at Total Mixture Dose[a]	Proportional Risk[a]
MCA	16	4.57E-04	11.93%	0	0
DCA	44.9	1.28E-03	33.48%	2.04E-05	6.8E-06
TCA	39.8	1.14E-03	29.68%	0	0
MBA	1.2	3.43E-05	0.89%	0	0
DBA[b]	0.8	2.29E-05	0.60%	2.04E-05	1.2E-07
BCA[b]	28.7	8.20E-04	21.40%	2.04E-05	4.4E-06
DCAN	1.6	4.57E-05	1.19%	0	0
TCAN	0.1	2.86E-06	0.07%	1.32E-04	9.8E-08
BCAN	0.7	2.00E-05	0.52%	1.01E-04	5.3E-07
DBAN	Not listed	—	—	—	—
DBCAN[c]	0.3	8.57E-06	0.22%	1.32E-04	2.9E-07
Sum	**—**	***3.83E-03***	***100.00%***	**—**	***1.2E-05***

[a]Based on dose-response data for total visceral malformations, using threshold dose-response model
[b]Estimated using the surrogate DCA
[c]Estimated using the surrogate TCAN

Table 5. Example Risk Estimate for Developmental Toxicity in Humans Exposed to Haloacetic Acids and Haloacetonitriles in Drinking Water at Concentrations Determined in Pilot Study, Ohio River, Following Chlorine Treatment and Simulated Distribution

Chemical	Water Concentration µg/L (95% UCL)	Estimated Dose mg/kg-day	Proportion of Component in Mixture	Risk at Total Mixture Dose[a]	Proportional Risk[a]
MCA	1.57	4.49E-05	2.12%	0	0
DCA	33.26	9.50E-04	44.80%	1.13E-05	5.0E-06
TCA	21.66	6.19E-04	29.17%	0	0
MBA	0.33	9.47E-06	0.45%	0	0
DBA[b]	1.66	4.75E-05	2.24%	1.13E-05	2.5E-07
BCA[b]	8.66	2.48E-04	11.67%	1.13E-05	1.3E-06
DCAN	4.23	1.21E-04	5.70%	0	0
TCAN	0.27	7.75E-06	0.37%	7.30E-05	2.7E-07
BCAN	2.32	6.64E-05	3.13%	5.57E-05	1.7E-06
DBAN[c]	0.27	7.72E-06	0.36%	7.30E-05	2.7E-07
DBCAN	Not listed		—	—	—
Sum	**—**	***2.12E-03***	***100.00%***	**—**	***8.9E-06***

[a]Based on dose-response data for total visceral malformations, using threshold dose-response model
[b]Estimated using the surrogate DCA
[c]Estimated using the surrogate TCAN

treatment produced lower concentrations of most of the haloacetic acids and haloacetonitriles, and lower total doses, but slightly higher concentrations of MCA, DBA, BCAN, and DBAN than did the parallel disinfection without ozone in the same pilot study. The estimated risk of developmental toxicity (based on dose-response for total visceral malformations) for the alternative ozone-chlorine treatment was 6.7×10^{-6}, slightly lower than that for the chlorine treatment shown in Table 5.

References

1. Hertzberg, R.C., G. Rice, L.K. Teuschler. 1999. Methods for Health Risk Assessment of Combustion Mixtures. In: Hazardous Waste Incineration: Evaluating the Human Health and Environmental Risks. S. Roberts, C. Teaf, and J. Bean eds. CRC Press LLC. p. 105-148

2. U.S. EPA. 1999. Guidance for Conducting Health Risk Assessment of Chemical Mixtures. External Review Draft. NCEA-C-0148.

3. U.S. EPA. 1986. Guidelines for the Health Risk Assessment of Chemical Mixtures. Federal Register. 51(185):34014-34025.

4. Teuschler, L.K. and R.C. Hertzberg. 1995. Current and future risk assessment guidelines, policy, and methods development for chemical mixtures. Toxicology 105, 137-144.

5. Jacangelo, J.G., N.L. Patania, K.M. Reagan, E.M. Aieta, S.W. Krasner and M.J. McGuire. 1989. Ozonation: Assessing its role in the formation and control of disinfection by-products. J. AWWA. 81:74-84.

6. Krasner, S.W., M.J. McGuire, J.G. Jacangelo, N.L. Patania, K.M. Reagan and E.M. Aieta. 1989. The occurrence of disinfection by-products in U.S. drinking water. J. Am. Water Works Assoc. 81:41-53.

7. Lykins Jr., B.W., W.E. Koffskey and K.S. Patterson. 1994. Alternative disinfectants for drinking water treatment. J. Environ. Eng. 120(4):745-758.

8. Miltner, R.J., E.W. Rice and A.A. Stevens. 1990. Pilot-Scale Investigation of the Formation and Control of Disinfection By-products. In: 1990 Annual Conference Proceedings. AWWA Annual Conference, Cincinnati, Ohio. 2:1787-1802.

9. Richardson, S.D. 1998. Identification of Drinking Water Disinfection By-Products. In: John Wiley's Encyclopedia of Environmental analysis and Remediation. Robert A. Meyers, Ed. Wiley and Sons. 3:1398-1421.

10. Bull, R.J. and F.C. Kopfler. 1991. Health Effects of Disinfectants and Disinfection By-Products. AWWA Research Foundation.

11. NTP. (National Toxicology Program) 1985. Toxicology and carcinogenesis studies of chlorodibromomethane in F344/N rats and B6C3F mice (gavage studies). NTP TR282.

12. NTP. (National Toxicology Program) 1986. Toxicology and carcinogenesis studies of bromodichloromethane in F344/N rats and B6C3F mice (gavage studies). NTP Technical Report, Ser. No. 321, NIH Publ. No. 87-2537.

13. NTP. (National Toxicology Program) 1989. Toxicology and carcinogenesis studies of tribromomethane and bromoform in F344/N rats and B6C3F mice (gavage studies). NTP-350.

14. Smith, M.K., J.L. Randall, E.J. Read and J.A.Stober. 1989. Teratogenic activity of trichloroacetic acid in the rat. Teratology 40: 445-451.

15. Piepel, G.F. and J.A. Cornell. 1985. Models for mixture experiments when the response depends on the total amount. Technometrics 27, 219-227.

16. Cornell, J.A. 1990. Experiments with Mixtures: Designs, Models, and the Analysis of Mixtures Data, Second Edition., John Wiley & Sons, Inc, New York, NY.

17. Chen, J.J., R.H. Heflich and B.S. Bass. 1989. A response-additive model for analyzing mixtures of two chemicals in salmonella reversion assay. Biomed. J. 31, 495-503.

18. Huckabee, J.W and N.A. Griffith. 1974. Toxicity of mercury and selenium to the eggs of carp (*Cyprinus carpio*). *Trans. Amer. Fish Soc*. 4, 822-825.

19. Birge, W.J., O.W. Roberts, and J.A. Black. 1976. Toxicity of metal mixtures in chick embryos. Bull. Environ. Contam. Toxicol. 16, 314-318.

20. Smith, M.K., J.L. Randall, E.J. Read and J.A.Stober. 1990. Developmental effects of chloroacetic acid in the Long-Evans rat. Teratology 41: 593.

21. Smith, M.K., J.L. Randall, E.J. Read and J.A.Stober. 1992. Developmental toxicity of dichloracetate in the rat. Teratology 46: 217-223.

22. Randall, J.L., S.A. Christ, P. Horton Perez, G.A. Nolen, E.J. Read and M.K. Smith. 1991. Developmental effects of 2-bromacetic acid in the Long-Evans rat. Teratology 43: 454.

23. Cicmanec, J.L., L.W. Condie, G.R. Olson and Shin-Ru Wang. 1991. 90-Day toxicity study of dichloroacetate in dogs. Fund. Appl. Toxicol. 17: 376-389.

24. Linder, R.E., G.R. Klinefelter, L.F. Strader, J.D. Suarez and N.L. Roberts. 1997. Spermatotoxicity of dichloroacetic acid. Reprod. Toxicol. 11: 681-688.

25. Linder, R.E., G.R. Klinefelter, L.F. Strader, J.D. Suarez, N.L. Roberts and C.J. Dyer. 1994. Spermatoxoicity of dibromacetic acid in rats after 14 daily exposure. Reprod. Toxicol. 8: 251-259.

26. Smith, M.K., J.L. Randall, J.A.Stober and E.J. Read. 1989. Developmental toxicity of dichloroacetonitrile: A by-product of drinking water disinfection. Fund. Appl. Toxicol. 12: 765-772.

27. Smith, M.K., J.L. Randall, D.R. Tocco, R.G.York, J.A.Stober and E.J. Read. 1988. Teratogenic effects of trichloroacetonitrile in the Long-Evans rat. Teratology. 38: 113-120.

28. Christ, S.A., E.J. Read, J.A. Stober and M.K. Smith. 1995. The developmental toxicity of bromochloroacetonitrile in pregnant Long-Evans rats. Int. J. Environ. Health Res. 5: 175-188.

OPTIMIZING THE PRECISION OF TOXICITY THRESHOLD ESTIMATION USING A TWO-STAGE EXPERIMENTAL DESIGN

Pamela F. Schwartz, Chris Gennings, Linda Teuschler and Marc W. Fariss[*]

* Pamela F. Schwartz is Senior Coordinator, Biometrics, Central Research Division, Pfizer Inc., Groton, CT 06340; Chris Gennings is Associate Professor, Department of Biostatistics, Virginia Commonwealth University, Medical College of Virginia, Richmond, VA 23298-0032; Linda Teuschler is a Mathematical Statistician at the U.S. Environmental Protection Agency, National Center for Environmental Assessment-Cincinnati Office, Cincinnati, Ohio, 45268; and Marc W. Fariss is Associate Professor, Department of Pharmaceutical Sciences, Washington State University, Pullman, WA 99164-6510.

August 1999

***Citation for this work: Schwartz, P.F., C. Gennings, L.K. Teuschler, M.W. Fariss. Submitted. Optimal Two-Stage Designs for a Threshold Model. Journal of Agricultural, Biological and Environmental Statistics.*

This work was supported by the U.S. Environmental Protection Agency's National Center for Environmental Assessment-Cincinnati Office in cooperative agreement CR 820847-01-0 while Pamela F. Schwartz was Research Associate, Department of Biostatistics, Virginia Commonwealth University, Medical College of Virginia, Richmond, VA and by a grant from the Virginia Center on Aging (MWF). The authors thank Xiao-Yan Gu, M.D. for the preparation of isolated hepatocytes and Sharon J. Fariss and Nancy C. Haar for technical assistance with the hepatocyte experiments.

Abstract

An important consideration for risk assessment is the existence of a threshold: the highest toxicant dose where the response is not distinguishable from background. We have developed methodology for finding an experimental design that optimizes the precision of threshold model parameter estimation for single-chemical threshold experiments. Being interested in precisely estimating the threshold parameter, we used the D-optimality and the D_s-optimality criteria. The D-optimal design results in parameter estimates as precise as possible in the sense that the likelihood-based confidence region has minimum volume, while the D_s-optimal threshold design results in parameter estimates as precise as possible in that the variance of the threshold parameter estimate is minimized. For nonlinear models, optimal designs are a function of the unknown parameters via the information matrix. Therefore, estimates of the parameters must be obtained before the optimal design of the experiment can be found. For this reason, a two-stage, D-D_s optimal design is recommended where the D-optimality criterion is used in the first stage followed by the D_s-optimality criterion in the second stage. The first stage is used for range finding and to obtain good global estimates to supply for the second stage. The second stage results in precise parameter estimates with minimum variance for the threshold parameter estimate. We propose that the use of this two-stage D-D_s optimal design will provide toxicologists with the experimental parameters necessary to accurately estimate thresholds for risk assessment purposes in a more cost effective and timely manner.

Introduction

In the toxicologic literature, the existence of thresholds has been debated for many years, and the use of a threshold concept as specifically applied in health risk assessment has been a controversial one (Crump, 1985; Daston, 1993; Hatch, 1971; Kodell et al., 1991; Mantel, 1963; Park and Snee, 1983). For both carcinogenic and noncancer effects, the assumption of a threshold depends on carefully defining the specific toxic effect of interest and then accurately estimating the threshold value. The definition of a threshold for risk assessment purposes revolves around three aspects of concern: whether the effect is clearly defined and *observable*, whether the effect is *expected* to occur and to what degree in humans, and whether the effect is not adverse and thus can be *ignored*. To estimate a threshold level from laboratory data, the effect must be quantifiable and "adverse" (e.g., not a reversible precursor effect, not a minor effect) such that the severity of the effect is of concern for humans. In addition, the extrapolation to humans from animal experiments is confounded because the effect may not be observable in the human population. This may be the result of the insensitivity of epidemiologic studies to detect such an effect, the general unavailability of relevant human data (the toxicity of the substance precludes the performance of human clinical studies), low human exposures, or physiological differences that allow effects to occur in animal studies that are not observed in humans.

Despite these concerns, the threshold concept is important to health risk assessment because of the need to protect sensitive subpopulations or sensitive individuals within a population. Thus, many of the chemical toxicity markers developed by risk assessors are estimates of threshold or subthreshold exposure levels, e.g., Reference Dose (RfD) and Reference Concentration (U.S. EPA, 1999), Acceptable Daily Intake (Lu, 1985), Maximum Tolerated Dose (Gaylor et al., 1985), and Tolerance Limit Value (Ulm, 1991). The RfD, for example, is typically calculated by dividing the No-Observed-Adverse-Effect Level (NOAEL) from an animal experiment by a number of appropriate uncertainty factors to estimate a lifetime exposure level that is not expected to produce deleterious effects in humans. Thus, the NOAEL, which by design is an experimental dose level, is critical to this calculation. An obvious improvement for this calculation is to replace the NOAEL with an estimation of the experimental threshold dose for a carefully defined and observable toxicologic endpoint of relevance to human health.

Biological evidence has supported the concept of a threshold (e.g., the importance of cellular repair and defense systems in preventing chemical-induced toxic responses, the dependence of embryotoxic responses on the reaction of groups of cells, tissues and organs rather than on single cells) and has thus stimulated further research (Ames et al., 1990; Gold et al., 1992; Haseman and Kupper, 1979; Kuchenhoff and Carroll, 1997; Rodricks et al., 1987; Wilson, 1973). Because of the importance of establishing regulatory "safe" levels for chemicals in the environment, it is essential that further work be done to define and estimate threshold values. Toward this end, this paper describes the development of a two-stage approach to estimating the lowest dose below which a defined toxic response (different from background) is

not observed (threshold dose) by specifying experimental designs that optimize the precision of threshold model parameter estimation for single-chemical experiments.

Many researchers have developed threshold models for single-chemical experiments (Thompson and Funderlic, 1981; Park and Snee, 1983; Crump, 1984; Cox, 1987, 1989; Ulm 1989, 1990, 1991; Kodell et al., 1991; Silvapulle, 1991; Schwartz et al., 1995a). Schwartz et al. (1995b) developed a multiple-chemical threshold model and methods for parameter estimation and confidence region construction. For a single chemical, the Schwartz et al. (1995a) threshold model is similar to Ulm's (1990) threshold model and can be parameterized as:

$$\mu_i = \begin{cases} \dfrac{1}{1+\exp\left(-\beta_0\right)}, & \text{if } d_i \le \delta \\ \dfrac{1}{1+\exp\left(-\left(\beta_o+\beta_1\left(d_i-\delta\right)\right)\right)}, & \text{if } d_i > \delta \end{cases} \tag{1}$$

where the ith dose level is represented by d_i (i=0, 1, ..., g), β_0 defines the background or spontaneous response, β_1 represents the slope of the dose-response curve, and δ represents the threshold dose value. The single-chemical threshold model defined in Equation (1) can be parameterized as:

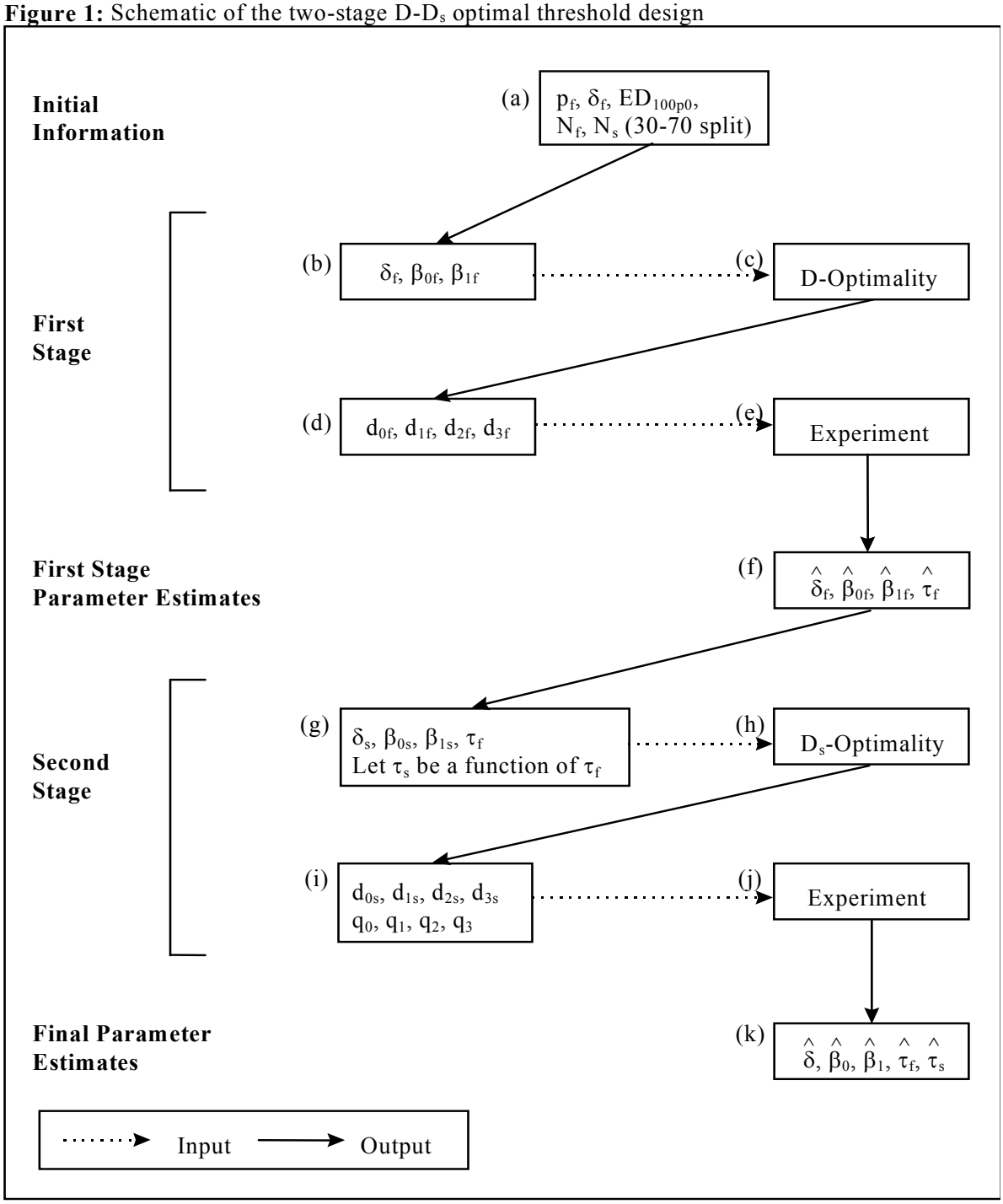
$$\mu_i = \begin{cases} \dfrac{1}{1+\exp\left(-\beta_0\right)}, & \text{if } \beta_1 d_i \le \delta^* \\ \dfrac{1}{1+\exp\left(-\left(\beta_o+\beta_1 d_i-\delta^*\right)\right)}, & \text{if } \beta_1 d_i > \delta^* \end{cases} \tag{2}$$

where $\delta^*=\beta_1\delta$. Notice the threshold model in Equation (2) is linear in the argument of the exponent; that is, $\beta_0+\beta_1 d_i-\delta^*$ is linear in the parameters. Both parameterizations of the threshold models given in Equations (1) and (2) are nonlinear because it is a piece-wise model.

Optimal designs for nonlinear models, and thus, threshold models, depend on the unknown parameters. For this reason, preliminary estimates of these unknown parameters are essential (Box and Lucas, 1959); thus, several researchers, including Abdelbasit and Plackett (1983), Minkin (1987) and Myers et al. (1995), suggest using a two-stage procedure that uses the second stage to complement the first.

A schematic of the two-stage D-D_s optimal threshold design using the quasi-likelihood methods detailed in this paper is shown in Figure 1 (The subscript "f" refers to the first stage and the subscript "s" refers to the second stage). The D-optimal design in the first stage produces global estimates for use in the second stage; the second stage uses D_s-optimality to generate precise parameter estimates with minimum variance for the threshold parameter estimate. First,

Figure 1: Schematic of the two-stage D-D_s optimal threshold design



NOTE: This diagram assumes equal allocation of observations among the dose groups in the first stage.

the biologic response of interest must be quantitatively defined as the percent response of interest (p_f) and a 100% response level (ED_{100p0}). An allocation is then made of the total number of experimental units to the first stage (N_f) and second stage (N_s) experiments. Initial information, such as results from a previous experiment, are used to make starting first stage parameter estimates for the threshold (d_f), background rate (b_{0f}), and slope (b_{1f}). D-optimality uses this initial information to determine the appropriate doses (d_{0f}, d_{1f}, d_{2f}, d_{3f}) for the first stage experiment. This process is then repeated using the first stage parameter estimates (including the first stage dispersion parameter, τ_f) as starting points for those of the second stage (d_s, b_{0s}, b_{1s} , τ_f). D_s-optimality uses these estimates to determine the appropriate doses (d_{0s}, d_{1s}, d_{2s}, d_{3s}), with respect to the proportion of experimental units in each dose group (q_0, q_1, q_2, q_3), for the second stage experiment. Using the results of the second experiment, final parameter estimates are made (including the second stage dispersion parameter, τ_s), yielding a final estimate of the threshold value.

Using similar results from the development of two-stage optimal designs for a logistic function, we develop two-stage D-D_s designs that optimize the precision of threshold model parameter estimation. We use quasi-likelihood methods to allow for dispersion. In Section 2 of this paper, we present two single-stage threshold design optimality criteria, namely D-optimality and D_s-optimality. In Section 3, we develop the two-stage D-D_s optimal threshold design using quasi-likelihood. In Section 4, we present an illustration of the actual two-stage D-D_s optimal threshold experiment using methods developed in Section 3. We conclude by discussing the flexibility of the procedure for conducting a two-stage D-D_s optimal threshold experiment. A discussion of relative efficiency of comparative designs is included in the appendix.

Single-Stage D-optimal and D_s-optimal Designs for a Threshold Model

Optimal design theory for linear regression has been well established in the literature (e.g. Kiefer and Wolfowitz, 1959). Some attention also has been given to optimal design theory for nonlinear regression, in particular, logistic regression (White, 1975; Kalish and Rosenberger, 1978; Abdelbasit and Plackett, 1983; Kalish, 1990; Chaloner and Larntz, 1989; Myers et al., 1994). For threshold modeling, we are primarily interested in parameter estimation and not prediction, and thus focus on D-optimality and D_s-optimality. We design experiments with three dose levels in addition to control because, in this way, we have a control dose and a low dose level to support estimation of the background response and two dose levels to support estimation of the slope.

Suppose there are g+1 (g=2, 3, ...) dose groups with m_i replicate observations for the ith dose. Let d_i represent the ith distinct dose group where d_0=0, and let $y_{ij} \in [0,1]$ denote the observed response for the jth replication in the ith dose group. Let m_i represent the number of experimental units receiving the ith dose level. Assuming the observed response, y_{ij}, for the jth replication in the ith dose group has mean μ_i defined in Equation (2) and variance $\mu_i(1-\mu_i)\tau$, the quasi-likelihood (Wedderburn, 1974; McCullagh, 1983; Liang and Zeger, 1986; Nelder and Pregibon, 1987; McCullagh and Nelder, 1989) allowing for dispersion (ignoring constants) is given by

$$Q(d^*, b_0, b_1, y_{ij}, d_i) = \frac{1}{t} SS\left\{ y_{if} \log\left(\frac{m_i}{1 - m_i} \right) + \log(1 - m_i) \right\} \quad (3)$$

Schwartz et al. (1995a) developed procedures for estimating threshold model parameters using quasi-likelihood methods. Closed-form solutions to the quasi-likelihood equations usually do not exist, and for threshold models, the form for the mean is not smooth everywhere. Thus, the Nelder-Mead simplex algorithm (Nelder and Mead, 1965; Olsson, 1974; Olsson and Nelson, 1975) is applied to maximize the quasi-likelihood to arrive at parameter estimates for the threshold model. The dispersion parameter, τ, is estimated using a moment estimator. Schwartz et al. (1995a) also developed methods for construction of a quasi-likelihood ratio confidence interval about the true threshold as well as a quasi-likelihood ratio statistic for testing for a significant dose-response relationship.

The D-optimal threshold design points are those design points that maximize with respect to **d** the determinant of the Fisher's quasi-information matrix; that is,

$$\underset{\mathbf{d}}{\textit{Maximum determinant of}} \frac{1}{\tau} I(\delta^*, \beta_0 \beta_1), \quad (4)$$

where $I(\delta^*, \beta_0, \beta_1)$ is the Fisher's quasi-information matrix (the negative of the matrix of second partial derivatives of $Q(\delta^*, \beta_0, \beta_1, y_{ij}, d_i)$ defined in Equation (3) with respect to δ^*, β_0, and β_1) and is evaluated at the specified values of δ^*, β_0, and β_1. The 4×1 vector of doses (3 dose levels in addition to control) is given by $\mathbf{d}=[d_0\ d_1\ d_2\ d_3]`$ where the number of dose levels is fixed, and τ is the dispersion parameter. Maximization of the determinant of the quasi-information matrix has the effect of minimizing the volume of the confidence ellipsoid based on an asymptotic normal approximation (Atkinson and Donev, 1992). The D-optimal threshold design results in parameter estimates as precise as possible in the sense that the likelihood-based confidence region has minimum volume.

D_s-optimality is similar to D-optimality except that the volume of the confidence ellipsoid for a subset of parameters is minimized (Atkinson and Donev, 1992). For the threshold model, we are primarily interested in precisely estimating the threshold, and thus, the D_s-optimality criterion is applied. The D_s-optimal threshold design points are those design points that minimize with respect to **d** the variance of the threshold parameter estimate; that is,

$$\underset{\mathbf{d}}{\textit{Minimum Var}}\ (\hat{\delta}^*), \quad (5)$$

where Var ($\hat{\delta}$ *) is the appropriate element in the inverse of I(δ^*, β_0, and β_1) and is evaluated at the specified values of δ^*, β_0, and β_1. The 4×1 vector of doses (3 dose levels in addition to control) is given by **d**=[d_0 d_1 d_2 d_3]` where the number of dose levels is fixed. The D_s-optimal threshold design results in parameter estimates as precise as possible in the sense that the variance of the threshold parameter estimate is minimized.

For the threshold model given in Equation (2), the D-optimal and D_s-optimal design criteria given in Equations (4) and (5), are a function of only dose, d_i, not response, y_{ij}. The logistic model that Minkin (1987) and Myers et al. (1995) considered can be linearized in terms of the logit. Although the threshold model defined in Equation (2) is linear in the parameters in the exponent, it is still a nonlinear model because it is a piece-wise model. The difference between the logistic model and the threshold model is important in how the optimization is carried out. Minkin (1987) and Myers et al. (1995) were able to derive optimal designs where optimization is with respect to the logit, whereas for threshold models, optimization is with respect to the doses, d_i.

Two-stage D-D_s Optimal Threshold Designs Using Quasi-Likelihood Methods

Similar to Abdelbasit and Plackett (1983), Minkin (1987), and Myers et al. (1995), we use a two-stage procedure that uses the second stage to complement the first stage because the optimal designs for a nonlinear model, and in particular a threshold model, depend on the unknown parameters. Myers et al. (1995) maintain that the first stage is used for range finding and to obtain good global estimates to supply to the second stage and that the second stage is the primary experiment of interest. We use a two-stage D-D_s optimal threshold design where the D-optimality criterion is used in the first stage followed by the D_s-optimality criterion in the second stage.

To determine the first stage D-optimal doses, the D-optimality criterion defined in Equation (4) is maximized with respect to **d**=[d_{0f} d_{1f} d_{2f} d_{3f}]`, a vector containing three dose levels in addition to control. I($\delta^*, \beta_0, \beta_1$) in Equation (4) is replaced with the first stage Fisher's quasi-information matrix, $(1/\tau_f)$ $I_f(\delta^*, \beta_0, \beta_1)$, and τ is replaced with the first stage dispersion parameter, τ_f. To determine the second stage D_s-optimal doses, the D_s-optimality criterion given in Equation (5) is minimized with respect to **d**=[d_{0s} d_{1s} d_{2s} d_{3s}]`, a vector containing three dose levels in addition to control, and with respect to q_i ($0 \le q_i \le 1$ such that $q_0+q_1+q_2+q_3=1$; i=0, 1, 2, 3), where q_i is the proportion of experimental units from the second stage allocated to d_{is}. *Var* ($\hat{\delta}$ *) in Equation (5) is replaced with the appropriate element in the inverse of the total conditional information for the two stages (Minkin, 1987), that is, the appropriate element in the inverse of

$$\{\frac{1}{\tau_f} I_f(\delta^*, \beta_0, \beta_1) + \frac{1}{\tau_s} I_s(\delta^*, \beta_0, \beta_1)\} \qquad (6)$$

The first and second stage Fisher's quasi-information matrices are given by $(1/\tau_f)I_f(\delta^*, \beta_0, \beta_1)$ and $(1/\tau_s)I_s(\delta^*, \beta_0, \beta_1)$, respectively, and τ_f and τ_s are the first and second stage dispersion parameters, respectively.

The D-optimal design in the first stage will result in good global estimates to supply to the second stage and the second stage will result in precise parameter estimates with minimum variance for the threshold parameter estimate. Similar to Myers et al. (1995), we allocate 30% of the total number of experimental units to the first stage and reserve 70% of the total number of experimental units for the second stage (30%-70% allocation). (The simulation study to be discussed in the appendix showed no differences between a 30%-70% allocation and 50%-50% allocation.) In this way, we are using fewer experimental units for range finding. Most of the experimental units are reserved for the second stage where we are ultimately interested in precise parameter estimates with minimum variance for the threshold parameter estimate. The final parameter estimates are obtained by using all the data combined.

Illustration

The administration of tacrine (tetrahydroaminoacridine), the only approved drug for the treatment of Alzheimer's disease, is hepatotoxic in nearly 50% of the patients who receive this drug (Watkins et al., 1994). To understand the molecular mechanism of tacrine-induced hepatotoxicity, Fariss et al. (1994) as well as other investigators are using isolated rat hepatocyte suspensions as an experimental model. To determine the mechanism of tacrine toxicity as well as potential species differences between rat and human (toxicity sensitivity and mechanism), we propose to estimate a threshold dose level where only doses above the threshold level result in hepatotoxicity. To determine tacrine-induced toxicity in rat hepatocytes, we use the previously published method of Fariss et al. (1994), and measure the percent of cellular lactate dehydrogenase (LDH) that leaks into the medium seven hours after treating with tacrine (mmol/L). Thus, the response of interest was the percent of LDH leakage seven hours after treatment.

As discussed in Section 1, optimal designs for nonlinear models, and thus threshold models, depend on the unknown parameters. For this reason, preliminary estimates of these unknown parameters are essential. Data from a previous experiment (Table 1) conducted in Fariss' laboratory provided the initial parameter estimates. Because all the LDH does not leak from the cell as the cell dies (presumably because cellular proteases released degrade the enzyme), the percentage of LDH leaked from the cell never reaches 100%. When this occurs (100% is not obtainable), the optimality criteria seek a dose with high percent LDH leakage. This high percent of LDH leakage may not be biologically possible and may be associated with a dose that exceeds the true region of activity. In other words, the high dose is meant to get a good slope estimate, not saturate the system. We adjusted each percent LDH leakage by dividing by the maximum observed percent LDH leakage and multiplying by 100.

Table 1. Initial Experimental Data

Dose (mmol/L)	Observed percent LDH leakage (%)	Adjusted* percent LDH leakage (%)
0.0	30	44.8
0.0	31	46.3
0.0	30	44.8
0.0	32	47.8
0.25	35	52.2
0.25	35	52.2
0.25	35	52.2
0.25	35	52.2
0.3	46	68.7
0.3	47	70.1
0.3	44	65.7
0.3	41	61.2
0.35	67	100.0
0.35	62	92.5
0.35	61	91.0
0.35	57	85.1
0.5	65	97.0
0.5	64	95.5
0.5	66	98.5
0.5	66	98.5

*Adjusted to obtain 100% LDH leakage by dividing by 67 and multiplying by 100.

For this experiment, the maximum observed percent LDH leakage was 67. Thus every percent LDH leakage was divided by 67 and multiplied by 100 (Table 1). Based on the adjusted observed percent LDH leakage, the initial parameter estimates were β_{0f}=-0.165, β_{1f}=17.407, and δ_f=0.237 (δ^*_f=4.125). We considered a design with three doses in the first stage (g_f=3) in addition to control and three doses in the second stage (g_s=3) in addition to control. In this way, for each stage, we have two doses to support estimation of the background and two doses to support estimation of the slope. The experimental system allows for between 15 and 20 flasks per experiment. Thus, N_f=16 flasks were allocated to the first stage with the flasks evenly distributed over the first stage doses; that is, $m_{0f}=m_{1f}=m_{2f}=m_{3f}$=4 ((a) and (b) in Figure 1), where m_{jk} is the number of flasks for the j^{th} control/dose group in the k^{th} stage experiment.

Based on the initial estimates β_{0f}, β_{1f}, and δ^*_f, and on the number of experimental units for the first stage, N_f, the D-optimality criterion in Equation (4) was applied, where $I(\delta^*, \beta_0, \beta_1)$ in Equation (4) was replaced with the first stage Fisher's quasi-information matrix, $(1/\tau_f)\, I_f(\delta^*, \beta_0, \beta_1)$, $\delta^*=\delta_f^*$, $B_0=B_{0F}$, $B_1 = B_{1f}$, and τ was replaced with the first stage dispersion parameter, τ_f. Because τ_f is a constant, $(1/\tau_f)\, I_f(\delta^*, \beta_0, \beta_1)$ reduces to $I_f(\delta^*, \beta_0, \beta_1)$ when maximization is performed ((c) in Figure 1). Given δ^*_f, β_{0f}, and β_{1f}, maximization was performed with respect to the first stage D-optimal doses, d_{1f}, d_{2f}, and d_{3f} where the control dose (d_{0f}=0) was included after maximization ((d) in Figure 1). For threshold models, it is necessary to apply the Nelder-Mead simplex algorithm to maximize the determinant of the quasi-information matrix. The first stage experiment was then conducted at d_{0f}=0, d_{1f}=0.162, d_{2f}=0.237, and d_{3f}=0.378 ((e) in Figure 1).

In an attempt to reach a dose that gives 100% LDH leakage, we added a dose level at d_{4f}=0.6. When the rat liver cells were harvested, there were only enough cells for 16 flasks. Thus, the experiment was conducted with N_f=15 and $m_{0f}=m_{1f}=m_{2f}=m_{3f}=m_{4f}$=3, where m_{jk} is the number of flasks for the j^{th} control/dose group in the k^{th} stage experiment. Table 2 shows the observed percent LDH leakage at the first stage D-optimal doses and at the additional dose (d_{4f}). As with the initial experiment, we divided the percent LDH leakage by 67. Based on the adjusted observed percent LDH leakage, y_{ij}, at d_{0f}=0, d_{1f}, d_{2f}, d_{3f}, and d_{4f}, the threshold model parameter estimates from the first stage were estimated as $\hat{\beta}_{0f}$=-0.708, $\hat{\beta}_{1f}$=1.238, and $\hat{\delta}_f$ =0.312 ($\hat{\delta}^*_f$ =0.386) using maximum quasi-likelihood estimation. The dispersion parameter, τ_f, was estimated as $\hat{\tau}_f$ =0.010 using a moment estimator (Schwartz et al., 1995a) ((f) in Figure 1).

For the second stage, the first stage parameter estimates were used as initial estimates; that is, $\delta_s = \hat{\delta}_f$, $\beta_{0s} = \hat{\beta}_{0f}$, and $\beta_{1s} = \hat{\beta}_{1f}$. Because we believed the experimental variability would be the same for the two stages, we assumed that $\tau_s = \hat{\tau}_f$ ((g) in Figure 1). For the second stage we assumed we had N_s=15 total flasks. Based on δ_s, β_{0s}, and β_{1s}, on $\hat{\tau}_f$ and τ_s, on the number of

Table 2. First Stage Experimental Data

Dose (mmol/L)	Observed percent LDH leakage	Adjusted* percent LDH leakage
0.0	24	35.8
0.0	21	31.3
0.0	21	31.3
0.162	21	31.3
0.162	21	31.3
0.162	22	32.8
0.237	21	31.3
0.237	22	32.8
0.237	26	38.8
0.378	23	34.3
0.378	24	35.8
0.378	23	34.3
0.6	36	53.7
0.6	24	35.8
0.6	23	34.3

*Adjusted to obtain 100% LDH leakage by dividing by 67 and multiplying by 100.

experimental units for the second stage, N_s, and on d_{0f}=0, d_{1f}, d_{2f}, d_{3f}, and d_{4f}, the D_s-optimality criterion given in Equation (5) was applied, where $Var(\hat{\delta}^*)$ in Equation (5) was replaced with the appropriate element in the inverse of the total conditional information for the two-stages given in Equation (6) ((h) in Figure 1). Given d_{0s}=0, δ_s, β_{0s}, β_{1s}, $\hat{\tau}_f$, τ_s, d_{0f}=0, d_{1f}, d_{2f}, d_{3f}, and d_{4f}, minimization was performed with respect to the second stage D_s-optimal doses, d_{1s}, d_{2s}, and d_{3s}, where the control dose was included before minimization. Minimization also was performed with respect to q_i ($0 \le q_i \le 1$ such that $q_0+q_1+q_2+q_3=1$; i=0, 1, 2, 3), where q_i is the proportion of experimental units from the second stage allocated to d_{is}. The number of experimental units allocated to the ith design point in the second stage is defined by $m_{is}=N_s q_i$ (i=0, 1, 2, 3) ((i) in Figure 1). For threshold models, it is necessary to apply the Nelder-Mead simplex algorithm to minimize the variance of the threshold parameter estimate.

The second stage experiment was then conducted at d_{0s}=0, (m_{0s}=2), d_{1s}=0.312 (m_{1s}=5), and d_{2s}=3.125 (m_{2s}=8) ((j) in Figure 1). The highest optimal dose, d_{2s}=3.125 may be too high for two reasons. First, no dose response was observed for the first stage, and second, we never achieved percent LDH leakage near 67. In other words, if d_{23} = 3.125 is too high and well beyond where the plateau begins in the dose-response curve, then the resulting slope estimate would be too shallow. Thus, we added a dose at 1.0 where we expected percent LDH would reach maximum response.

Table 3 shows the observed percent LDH leakage at the first and second stage D-optimal and D_s-optimal doses, respectively, in addition to the extra doses at 0.6 and 1.0. As with the initial experiment and the first stage experiment, we wanted to divide the percent LDH leakage by the maximum percent LDH leakage, but as it turned out the observed maximum (90%) was greater than expected (67%). Thus, we divided by 90 (and multiplied by 100) the maximum observed percent LDH leakage for the second stage.

Based on the adjusted observed percent LDH leakage, y_{ij}, at d_{0f}=0, d_{1f}, d_{2f}, d_{3f}, d_{0s}=0, d_{1s}, d_{2s}, and d_{3s}, the final threshold model parameters were estimated as $\hat{\beta}_0$ =-1.077, $\hat{\beta}_1$ =2.168, and $\hat{\delta}$ =0.939 ($\hat{\delta}^*$ =2.036) using maximum quasi-likelihood estimation, and τ_s was estimated as $\hat{\tau}_s$ =0.011 using moment estimation (Schwartz et al., 1995a) ((k) in Figure 1). The estimated 95% Wald confidence interval about the threshold is given by (0.832, 1.045), indicating that a significant threshold was detected. The lower limit of the 95% Wald confidence interval, 0.832 mmol/L of tacrine, is a conservative estimate of the dose where the rate of hepatotoxicity begins to increase above background.

To validate of the final estimate of the threshold dose-response relationship, an experiment was conducted at the control dose, 0.0, at a dose below the lower limit on the estimated 95% Wald confidence interval about the threshold, 0.75, at a dose at the estimated threshold, 0.94, and at a dose above the upper limit on the estimated 95% Wald confidence

Table 3. First and Second Stage Experimental Data

Dose (mmol/L)	Observed percent LDH leakage (%)	Adjusted* percent LDH leakage (%)
0.0	24	26.7
0.0	21	23.3
0.0	21	23.3
0.162	21	23.3
0.162	21	23.3
0.162	22	24.4
0.237	21	23.3
0.237	22	24.4
0.237	26	28.9
0.378	23	25.6
0.378	24	26.7
0.378	23	25.6
0.6	36	40.0
0.6	24	26.7
0.6	23	25.6
0.0	21	23.3
0.0	22	24.4
0.312	22	24.4
0.312	22	24.4
0.312	21	23.3
0.312	22	24.4
0.312	21	23.3
1.0	37	41.1
1.0	20	22.2
1.0	20	22.2
1.0	23	25.6
1.0	26	28.9
3.125	86	95.6
3.125	88	97.8
3.125	87	96.7
3.125	86	95.6
3.125	89	98.9
3.125	90	100.0
3.125	87	96.7
3.125	89	98.9

*Adjusted to obtain 100% LDH leakage by dividing by 90 and multiplying by 100.

interval about the threshold, 1.5. As expected (Table 4), at the control dose and the dose below the estimated threshold, we did not obtain percent LDH leakage above background, whereas at the estimated threshold and at a dose above the estimated threshold, we did obtain an increase in percent LDH leakage.

The experiment used for initial information was conducted by a different investigator from the first, second, and validation stages, which were conducted by Fariss. The cell densities from the first, second, and validation stages were consistent, but were different (greater) from the cell densities from the initial information. We assume that the initial information did not hit the region of activity because of differences in cell density and in experimental protocol. In the initial information experiments only, multiple samples were removed from, the suspensions every hour for seven hours thus significantly reducing suspension volume. The dose response after the first stage was significantly affected by a single response at 0.6 mmol/L of tacrine. Because we missed the region of activity, the dose response after the first stage is more shallow than expected. After the second stage, the dose-response relationship is as expected. Figure 2 shows the observed percent LDH leakage as well as estimated threshold curves. To plot the results from the experiment, we divided the percent LDH leakage in stage 1, stage 2, and the validation stage by 90. Next, we adjusted the initial curve and the first stage estimated threshold curve by multiplying by 0.74 (67/90). Notice that the initial information did not give percent LDH above background.

Conclusions

To determine the two-stage D-D_s optimal threshold design, the investigator must supply, in addition to preliminary estimates for the unknown parameters and the total sample size for the first and second stages (N_f and N_s), the number of replications per experimental unit for the first and second stages, namely n_{ijf} and n_{ijs}. The number of replications per experimental unit can be based on past experience, or in the special case of reproductive or developmental data, can be randomly generated from the table of litter sizes given by Kupper et al. (1987).

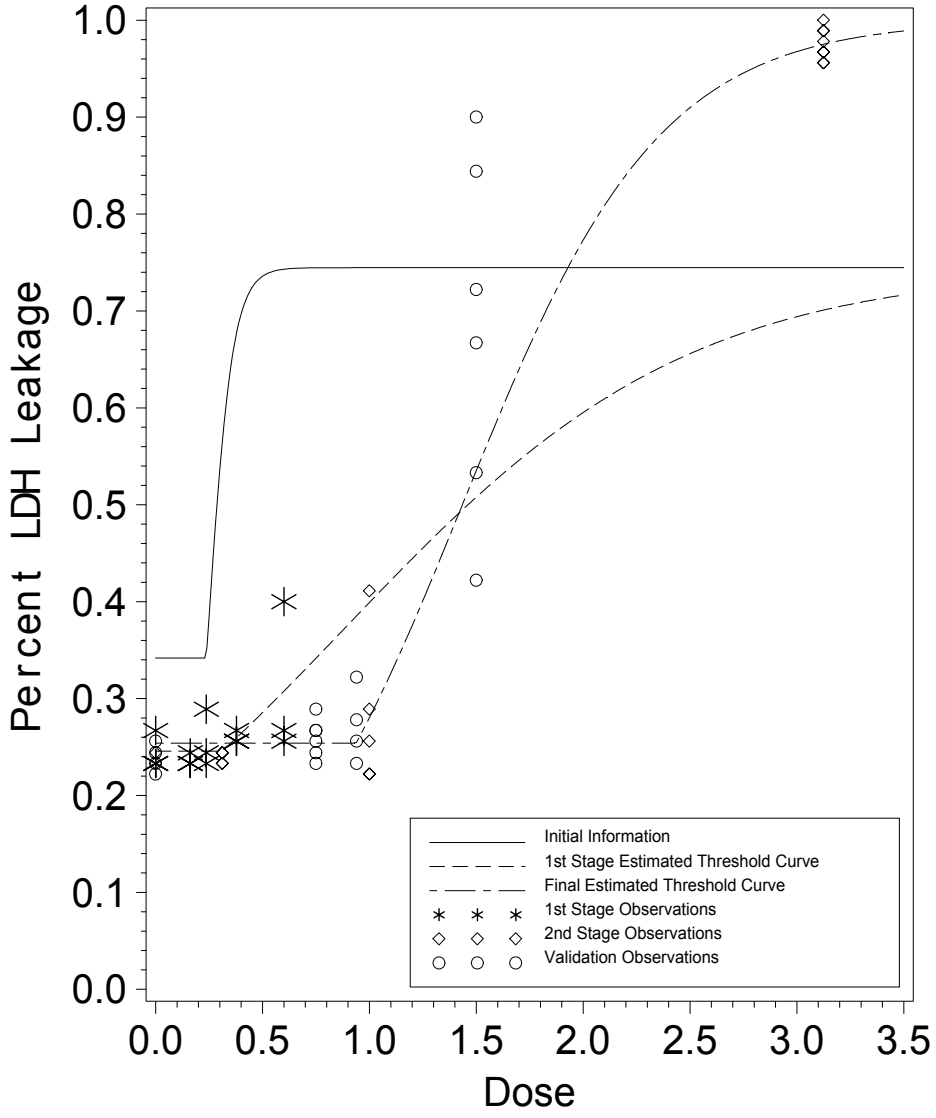
The methods for deriving two-stage D-D_s optimal threshold designs described above can be very flexible. Because of this flexibility, the toxicologist and statistician should collaborate prior to designing an experiment to determine the most efficient yet practical threshold design. Based on this collaboration, several areas of the design process can be modified to meet individual needs. First, unequal optimal allocation of the experimental units to the first stage design points may be used. Second, based on the agreement of the initial estimates and the resulting parameter estimates from the first stage, a decision may be made to stop the experiment after the first stage. Third, to ensure at least one design point, other than control, falls below the initial estimate of the threshold, it may be necessary to include a control dose only after the D-optimal design points have been determined for the first stage. Based on our experience with toxicologists, their intuition says to put more design points close to the threshold. On the other hand, based on our experience and a simulation study (Schwartz, 1992), assuming the model in Equation (2) is appropriate, it is more important to have good estimates for β_0 and β_1. From

Table 4. Validation Stage Experimental Data

Dose(mmol/L)	Observed percent LDH leakage (%)	Adjusted* percent LDH leakage (%)
0.0	20	22.2
0.0	22	24.4
0.0	23	25.6
0.0	21	23.3
0.75	24	26.7
0.75	26	28.9
0.75	21	23.3
0.75	22	24.4
0.75	23	25.6
0.75	24	26.7
0.94	25	27.8
0.94	29	32.2
0.94	21	23.3
0.94	23	25.6
1.5	38	42.2
1.5	65	72.2
1.5	81	90.0
1.5	76	84.4
1.5	48	53.3
1.5	60	66.7

*Adjusted to obtain 100% LDH leakage by dividing by 90 and multiplying by 100.

Figure 2
Observed Data and Threshold Dose-Response Curves



these, a good estimate of the threshold parameter follows as a result of continuity of the threshold model defined in Equation (2). Thus, it is important to place design points that result in a precise estimate of the background and the slope. A precise estimate of the slope generally follows from a design that spreads the design points out. For example, in a linear model, the variance of the slope estimate is minimized by using a two-point design at the extremes of the experimental region, i.e., as far away as possible from the center of the design. As in the illustration, a toxicologist may decide to add several design points in a validation (third) stage around the estimate of the threshold, $\hat{\delta}_f^*$ to confirm the final results. Fourth, a different dispersion parameter can be estimated for each stage, that is, τ_f and τ_s. Differences in the dispersion parameter estimates, $\hat{\tau}_f$ and $\hat{\tau}_s$, may reveal differences in experimental conditions between the two stages.

In conclusion we recommend the use of our two-stage, D-D_s optimal design for risk assessment to provide toxicologists with a cost effective and time-efficient method for determining toxicity threshold estimates. Additional experimental trials using this design will be required to further validate its effectiveness in threshold estimation and risk assessment.

Appendix

We have described both single and two-stage designs for 'optimal' estimation of threshold parameters. In planning a study, investigators must choose the criterion over which they desire an optimum design. Therefore, methods for comparing these designs and their subsequent properties are discussed here. The decision to use either a single stage or two-stage design may be based on time constraints and/or the perceived success of the first stage in describing the dose-response relationship.

Comparison of Single-Stage D-Optimal and D_s-optimal Threshold Designs

Relative D-efficiency or D_s-efficiency can be used to compare the single-stage D-optimal and the single-stage D_s-optimal threshold designs. The D-efficiency of the single-stage D_s-optimal threshold design relative to the single-stage D-optimal threshold design is defined by

$$D-eff+= \{\frac{(Maximum\ determinant\ of\ I(\delta^*,\beta_0,\beta_1)\ for\ the\ D-optimal\ threshold\ design)IN}{(Maximum\ determinant\ of\ I(\delta^*,\beta_0,\beta_1)\ for\ the\ D_s-optimal\ threshold\ design)IN}\}^{1/p},$$

and the D_s-efficiency of the single-stage D-optimal threshold design relative to the single-stage D_s-optimal threshold design is defined by

$$D_s-eff+= \{\frac{(MinimumVar(\delta^*)\ for\ the\ D-optimal\ threshold\ design)IN}{(MinimumVar(\delta^*)\ for\ the\ D_s-optimal\ threshold\ design)IN}\}^{1/p},$$

where N is the total number of experimental units and p is the number of parameters. For threshold models, D-efficiency and D_s-efficiency are functions of the unknown parameters, δ^*, β_0, and β_1.

As a comparison of the single-stage D-optimal and single-stage D_s-optimal threshold designs, a small simulation study was performed (Schwartz et al., 1995b) where D-efficiencies and D_s-efficiencies are calculated for 24 unique cases for the slope, background, and threshold parameters. The two levels for the background are $\beta_0 = -3$ and $\beta_0 = -2$, resulting in a background response of 0.05 and 0.12, respectively. The four levels for the slope are $\beta_1 = 2.5$, $\beta_1 = 5.0$, $\beta_1 = 7.5$, $\beta_1 = 10.0$ with $\beta_1 = 2.5$, resulting in a rather "shallow " slope, and $\beta_1 = 10.0$, resulting in a rather "steep" slope. The three levels for the threshold are $\delta = 0.2$, $\delta = 0.3$, and $\delta = 0.4$ where we expect the doses to range between 0 and 3. The results of the simulation study indicate that the single-stage D-optimal threshold designs are more D-efficient than the single-stage D_s-optimal threshold, and that the single-stage D_s-optimal threshold designs are more D_s-efficient than the single-stage D-optimal threshold designs. The results also indicate that for both the single-stage D-optimality and D_s-optimality criteria, a smaller background (β_0), or a more "shallow" slope

(β_1), or a larger threshold (δ^*), tend to result in higher optimal doses. Generally, the D_s-optimal doses tend to result in a wider range of responses. Unlike several of the optimal designs for the logistic regression model (White, 1975; Kalish and Rosenberger, 1978; Abdelbasit and Plackett, 1983), the single-stage D_s-optimal and D-optimal threshold designs are not symmetric.

Two-Stage Optimality Criteria and Conditional Relative Efficiency

To compare the two-stage D-D_s optimal threshold design to the two-stage D-D optimal threshold design, we must determine the two-stage D-optimality and D_s-optimality criteria. This is not as straightforward as it may seem because the second stage is conditional on the first stage results, and thus, the second stage doses and sample allocations (d_{0s}, d_{1s}, d_{2s}, d_{3s}, m_{0s}, m_{1s}, m_{2s}, and m_{3s}) are random variables dependent on the parameter estimates obtained in the first stage. Thus, the two-stage D-optimality and D_s-optimality criteria must account for this randomness. From Myers et al. (1995) and the basic laws of probability, we know that

$$\mathrm{Var}(\mathbf{b}) = \underset{\mathbf{t}}{\mathrm{E}}\{\mathrm{Var}(\mathbf{b}|\mathbf{t})\} + \underset{\mathbf{t}}{\mathrm{Var}}\{\mathrm{E}(\mathbf{b}|\mathbf{t})\} \tag{7}$$

where the vector **t** represents the random variables (d_{0s}, d_{1s}, d_{2s}, d_{3s}, m_{0s}, m_{1s}, m_{2s}, m_{3s}) and the first stage parameter estimates. The vector **b** represents the estimator of the vector of model parameters for the two-stage procedure. When calculating the Fisher's quasi-information matrix, or the asymptotic variance-covariance matrix for the two-stage procedure, and thus the two-stage D-optimality and D_s-optimality criteria, a simulation study (Schwartz et al., 1995a) is necessary to calculate the empirical expectation in the first term in Equation (7). The second term in Equation (7) is asymptotically zero. Thus, the D_s-optimality and D-optimality criteria are obtained by averaging over the simulated replications.

Conditional relative D-efficiency and conditional relative D_s-efficiency (Atkinson and Donev, 1992) are useful tools for comparing designs. For a two-stage design, the conditional D-efficiency and D_s-efficiency of design "A" relative to design "B" are defined as

$$Conditional\,relative\,D-efficiency = \left(\frac{E_B}{E_A}\right)^{1/p}$$

and

$$\mathrm{Conditional\,relative\,D_s-efficiency} = \left(\frac{E_A}{E_B}\right)^{1/p}$$

respectively, where

$$E_A = \frac{\text{Two-stage D-optimality of } D_s\text{-optimality criterion for design "A"}}{N_A}$$

and

$$E_B = \frac{\text{Two-stage D-optimality of } D_s\text{-optimality criterion for design "B"}}{N_B}$$

are obtained by simulation (Schwartz et al., 1995a). The total number of experimental units for design "A" is given by N_A; the total number of experimental units for design "B" is given by N_B; and p is the number of parameters in the model.

Notice the conditional relative D-efficiency and the conditional relative D_s-efficiency are adjusted for the number of experimental units, and thus, models with different numbers of experimental units are scaled for comparison. If the conditional D-efficiency of design "A" relative to design "B" is greater than one, design "B" is a more D-efficient design. For design "B", each individual experimental unit is contributing more information to the final parameter estimates than each individual experimental unit is contributing for design "A." If the conditional D_s-efficiency of design "A" relative to design "B" is greater than one, design "B" is a more D_s-efficient design.

Properties of the Two-Stage D-D_s Optimal Threshold Design

To compare the two-stage D- D_s optimal and D-D optimal threshold designs we can look at the two-stage conditional relative D-efficiency and the two-stage conditional relative D_s-efficiency. The two-stage optimality criteria are conditional on the first stage optimal doses, parameter estimates, and sample size. Thus, to determine the two-stage relative efficiencies, we must first determine the two-stage D_s-optimality and D-optimality criteria via a simulation study. The simulation for a two-stage D- D_s optimal threshold design and the two-stage D-D optimal threshold design was performed as follows:

1. Select values for the true, unknown parameters ($\beta_0 = -2.0$, $\beta_1 = 12$, and $\delta = 0.2$).

2. Select a total sample size ($N = N_f + N_s$) and the proportion of data for the first and second stage. Two total sample sizes were selected (N = 80 and N = 160) with either 50% of the data in the first stage and 50% percent in the second stage (50%-50% allocation) or 30% of the data in the first stage and 70% percent in the second stage (30%-70% allocation). Thus, four different sampling situations

were simulated (a) $N_f = 40$ and $N_s = 40$, (b) $N_f = 24$ and $N_s = 56$, (c) $N_f = 80$ and $N_s = 80$, and (d) $N_f = 48$ and $N_s = 112$.

3. Select values for the initial (erroneous) parameter estimates. These are the values supplied by the investigator prior to designing the experiment. Three values were chosen for the background parameter ($\beta_{0f} = -2.5, -2.0$, and -1.5); three values were chosen for the slope parameter ($\beta_{1f} = 6, 9$, and 12); and three values were chosen for the threshold parameter ($\delta_f = 0.1, 0.2$, and 0.3). Thus, 27 cases for the erroneous initial parameter estimates were simulated.

4. For given sample sizes ((2) above) and erroneous initial parameter estimates ((3) above), the first stage D-optimal design is obtained as illustrated above in this Appendix.

5. The second stage D_s-optimal design points are obtained as illustrated above in this Appendix, and the D-optimal design points are obtained similarly.

6. The second stage D_s-optimal design points in (5) above are replicated for 100 randomly generated beta-binomial data sets. For each replication, the D_s-optimality and the D-optimality criteria are calculated for the D_s-optimal design. The second stage D-optimal design points in (5) above are also replicated for 100 randomly generated beta-binomial data sets. For each replication, the D_s-optimality and the D-optimality criteria are calculated for the D-optimal design.

7. The D_s-optimality criterion and the D-optimality criterion are averaged across the 100 replications to obtain the two-stage D_s-optimality and D-optimality criteria for the two-stage D- D_s optimal threshold design and the two-stage D-D optimal threshold design. The two-stage D-optimality criterion is given by

Two-stage D-optimality = Mean(D-optimality criterion),

and the two-stage D_s-optimality criterion is given by

Two-stage D_s-optimality = Mean(D_s-optimality criterion).

The main conclusion from this simulation study is that regardless of the total number of experimental units, the first and second stage sample allocations, and the erroneous initial parameter estimates, there are no significant differences between the two-stage D-D_s optimal threshold design and the two-stage D-D optimal threshold design with respect to conditional relative D_s-efficiency, the two-stage D_s-optimality criterion, the percent of times at least two doses are less than the estimate of the threshold after the first stage ($\hat{\delta}_f^*$), or the percent of times there are four unique doses (including control) in the second stage. Based on our experience, toxicologists are more comfortable with threshold estimates derived from several doses below

the threshold. Thus, to convince a toxicologist that the final parameter estimates are reasonable, it is important to have at least one design point, other than control, below the first stage estimate of the threshold. Generally, both designs have four unique optimal doses (including control) with at least one experimental unit allocated to each optimal dose.

References

Abdelbasit, K. M. and R. L. Plackett. 1983. Experimental design for binary data. *Journal of the American Statistical Association* **78**, 90-98.

Ames, B. N., M. Profet, and L. S. Gold. 1990. Nature's chemicals and synthetic chemicals: Comparative toxicology. *Proceedings of the National Academy of Science USA* **87**, 7782-7786.

Atkinson, A. and A. Donev. 1992. *Optimum Experimental Designs*. Oxford: Oxford University Press.

Box, G. E. P. and H. L. Lucas. 1959. Design of experiments in non-linear situations. *Biometrika* **46**, 77-90.

Chaloner, K. and K. Larntz. 1989. Optimal bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference* **21**, 191-208.

Cox, C. 1987. Threshold dose-response models in toxicology. *Biometrics* **43**, 511-523.

Cox, C. 1989. Correspondence. *Biometrics* **45**, 1327-1328.

Crump, K. S. 1984. A new method for determining allowable daily intakes. *Fundamental and Applied Toxicology* **4**, 854-871.

Crump, K. S. 1985. Methods for carcinogenic risk assessment. In *Principles of Health Risk Assessment*, P. F. Ricci (ed.), 279-319. New Jersey: Prentice-Hall.

Daston, G. P. 1993. Do thresholds exist for developmental toxicants? In *Issues and Reviews in Teratology, Volume 6*, H. Kalter (ed.), New York: Plenum Press.

Fariss, M. W., S. A. Johnsen, L. P. Walton, V. R. Mumaw, and S. D. Ray. 1994. Tetrahydroaminoacridine-induced ribosomal changes and inhibition of protein synthesis in rat hepatocyte suspensions. *Hepatology* **20**, 240-246.

Gaylor, D. W., J. J. Chen, and R. L. Kodell. 1985. Experimental design of bioassays for screening and low dose extrapolation. *Risk Analysis*. **5**, 9-16.

Gold, L. S., T. H. Slone, B. R. Stern, N. B. Manley, and B. N. Ames. 1992. Rodent carcinogens: Setting priorities. *Science* **258**, 261-265.

Haseman, J. K. and L. L. Kupper. 1979. Analysis of dichotomous response data from certain toxicological experiments. *Biometrics* **35**, 281-292.

Hatch, T. F. 1971. Thresholds: Do they exist? *Archives of Environmental Health* **22**, 687-689.

Kalish, L. A. 1990. Efficient designs for estimation of median lethal dose and quantal dose-response curves. *Biometrics* **46**, 737-748.

Kalish, L. A. and J. L. Rosenberger. 1978. Optimal designs for the estimation of the logistic function. *Technical Report 33*, Pennsylvania State University.

Kiefer, J. and J. Wolfowitz. 1959. Optimum designs in regression problems. *The Annals of Mathematical Statistics* **30**, 271-294.

Kodell, R. L., R. B. Howe, J. J. Chen, and D. W. Gaylor. 1991. Mathematical modeling of reproductive and developmental toxic effects for quantitative risk assessment. *Risk Analysis* **11**, 583-590.

Kuchenhoff, H. and R.J. Carroll. 1997. Segmented regression with errors in predictors: Semi-parametric and parametric methods. *Statistics in Medicine*. **16**, 169-188.

Kupper, L. L., C. Portier, M. D. Hogan, and E. Yamamoto. 1986. The impact of litter effects on dose-response modeling in teratology. *Biometrics* **42**, 85-98.

Liang, K. and S. Zeger. 1986. Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 12-22.

Lu, F.C. 1985. Safety assessments of chemicals with threshold effects. *Regulatory Toxicology and Pharmacology*. **5**, 460-464.

Mantel, N. 1963. The concept of threshold in carcinogenesis. *Clinical Pharmacology and Therapeutics* **4**, 104-109.

McCullagh, P. 1983. Quasi-likelihood functions. *The Annals of Statistics* **11**, 59-67.

McCullagh, P. and J. A. Nelder. 1989. *Generalized Linear Models*, Second Edition. New York: Chapman and Hall.

Minkin, S. 1987. Optimal designs for binary data. *Journal of the American Statistical Association* **82**, 1098-1103.

Myers, W. R., R. H. Myers, and W. H. Carter, Jr. 1994. Some alphabetic optimal designs for the logistic regression model. *Journal of Statistical Planning and Inference* **40**, 57-77.

Myers, W. R., R. H. Myers, W. H. Carter, Jr., and K. L. White, Jr. 1995. Two-stage designs for the logistic regression model in single agent bioassays. *Communication in Statistics* (In review).

Nelder, J. A. and R. Mead. 1965. A simplex method for function minimization. *The Computer Journal* **7**, 308-313.

Nelder, J. A. and D. Pregibon. 1987. An extended quasi-likelihood function. *Biometrika* **74**, 221-232.

Olsson, D. M. 1974. A sequential simplex program for solving minimization problems. *Journal of Quality Technology* **6**, 53-57.

Olsson, D. M. and L. S. Nelson. 1975. The Nelder-Mead simplex procedure for function minimization. *Technometrics* **17**, 45-51.

Park, C. N. and R. D. Snee. 1983. Quantitative risk assessment: State-of-the-art for carcinogenesis. *The American Statistician* **37**, 427-441.

Rodricks, J. V., V. Frankos, D. Turnbull, and R. G. Tardiff. 1987. Risk assessment for effects other than cancer. In *Food Protection Technology*, Charles W. Felix (ed.), 61-74. Chelsea, MI: Lewis.

Schwartz, P. F. 1992. *Threshold Models in Risk Assessment for a Combination of Agents*. Unpublished Ph.D. thesis, Richmond, VA: Virginia Commonwealth University, Department of Biostatistics.

Schwartz, P. F., C. Gennings, and M. W. Fariss. 1995a. Optimal Designs for a Threshold Model: Simulation Results. Technical Report #9501. Richmond, VA: Virginia Commonwealth University, Department of Biostatistics.

Schwartz, P. F., C. Gennings, and V. M. Chinchilli. 1995b. Threshold Models for Combination Data from Reproductive and Developmental Experiments. *Journal of the American Statistical Association* **90**(431), 862-870.

Silvapulle, M. J. 1991. On testing for threshold values. *Biometrics* **47**, 1628-1629.

Thompson, W. A. and R.E. Funderlic. 1981. A simple threshold model for the classical bioassay problem. In *Measurements of Risks*, G. G. Berg and H. D. Maillie (ed.), New York: Plenum Press, p. 521-533.

Ulm, K. W. 1989. On the estimation of threshold values. *Biometrics* **45**, 1324-1326.

Ulm, K. W. 1990. Threshold models in occupational epidemiology. *Mathematical and Computer Modelling* **14**, 649-652.

Ulm, K. W. 1991. A statistical method for assessing a threshold in epidemiological studies. *Statistics in Medicine* **10**, 341-349.

U.S. EPA. 1999. Integrated Risk Information System (IRIS). Online. National Center for Environmental Assessment, Washington, DC.

Watkins, P. B., H. J. Zimmerman, M. J. Knapp, S. I. Gracon, and K. W. Lewis. 1994. *Journal of the American Medical Association* **271**, 992-998.

Wedderburn, R. W. M. 1974. Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* **61**, 439-447.

White, L. V. 1975. *The Optimal Design of Experiments for Estimation in Nonlinear Models*. Unpublished Ph.D. thesis. Department of Mathematics, Imperial College, London.

Wilson, J. G. 1973. *Environment and Birth Defects*. New York: Academic Press.

APPENDIX A

DBP MIXTURES RESEARCH CITATIONS

PUBLICATIONS/REPORTS

Gennings, C., P. Schwartz, W.H. Carter, Jr., J.E. Simmons. 1997. Detection of departures from additivity in mixtures of many chemicals with a threshold model. J. Agric. Biol. Environ. Stat. 2:198-211. Erratum. 2000. J. Agric. Biol. Environ. Stat. 5(2):257.

Gennings, C., L. Teuschler, W. Hartley, A. Thiyagarajah, J.E. Simmons. 1999. Novel Statistical Methods for Risk Assessment of Water Disinfection By-products. AWWA 1999 Annual Conference Proceedings. Chicago, IL. TU 22.2 (CD ROM).

Hartley, W.R., C. Gennings, L. Teuschler, A. Thiyagarajah, J.E. Simmons. 1999. Advances in the Toxicological Assessment of Disinfection By-Products in Rodent and Fish Biomedical Models. AWWA 1999 Annual Conference Proceedings. Chicago, IL. TU 22.1 (CD ROM).

Schwartz, P.F., C. Gennings, L.K. Teuschler, M.W. Fariss. 2000. Optimal two-stage designs for a threshold model. J. Agric. Biol. Environ. Stat. (In Press)

Simmons, J.E., L.K. Teuschler, C. Gennings. 2000. The Toxicology of DBP Mixtures: Methods for Multi-Chemical Assessment, Present Research Efforts, and Future Research Directions. In: Microbial Pathogens and Disinfection By-Products in Drinking Water: Health Effects and Management of Risks, G.F. Craun, F.S. Hauchman, D.E. Robinson, Ed. ILSI Press. (In Press)

Stiteler, W.M., L.K. Teuschler, J.T. Colman, R.C. Hertzberg. 2000. A Proportional Response Additivity Model for Evaluating Health Risks Associated with Complex Mixtures: An Example using Drinking Water Disinfection By Products. (Submitted to Risk Analysis)

Syracuse Research Corporation. 1998. Final Report. Methodology for Improving Risk Assessment. EPA Cooperative Agreement #CR-822761-01. October.

Teuschler, L.K., C. Gennings, W.M. Stiteler, R.C. Hertzberg, J.T. Colman, A. Thiyagarajah, J.C. Lipscomb, W.R. Hartley, J.E. Simmons. 2000. A multiple-purpose design approach to the evaluation of risks from complex mixtures of disinfection by-products (DBPs). Drug Chem. Toxicol. 23(1):307-321. Erratum. Accepted for Publication in 2000. Drug Chem. Toxicol.

Tulane University Medical Center. 1999. Final Report. The Use of the Japanese Medaka Fish (*Oryzias latipes*) for the Evaluation of the Carcinogenic Potential and Developmental Effects of Mixtures of Contaminants. Cooperative Agreement #CR-822766-01. August.

U.S. EPA. 1998. Workshop Report: The Risk Assessment of Mixtures of Disinfection By-Products (DBPs) in Drinking Water. NCEA-C-0459.

Virginia Commonwealth University. 1995. Final Report. Design Considerations for Threshold Models in Risk Assessment. Cooperative Agreement #820847-01. May.

Virginia Commonwealth University. 1998. Final Report. Chemical Mixtures Risk Assessment. Cooperative Agreement #CR822671-01. October.

POSTER PRESENTATIONS

Hartley, W.R., A. Thiyagarajah, L. Teuschler, C. Gennings, C. Cubbison. 1999. Exposure of Japanese Medaka (*Oryzias Latipes*) Embryos to Benzene, Chloroform, and Bromoform for Assessment of Carcinogenic Potential. Society of Toxicology 38th Annual Meeting. March.

Hartley, W.R., J.C. Lipscomb, L.K. Teuschler, C. Gennings, A. Thiyagarajah. 2000. Chronic Toxicity of Chloroform in the Japanese Medaka (Oryzias Latipes). Society of Toxicology 39th Annual Meeting, Philadelphia, PA. March.

McDonald, A., C. Gennings, Y.M. Sey, W.H. Carter, Jr., L.K. Teuschler, J.E. Simmons. 1997. Hepatotoxicity of Trihalomethanes (THMs) in Female Mice. Society of Toxicology 36th Annual Meeting. March.

McDonald, A., C. Gennings, W. Hartley, L.K. Teuschler, A. Thiyagarajah, Y.M. Sey, J.E. Simmons. 1999. The Effect of Mixing Ratio on Mixtures of Trihalomethane Disinfection By-products. Society of Toxicology 38th Annual Meeting. March.

Schwartz, P., C. Gennings, M. Fariss, L. Knauf, R. Hertzberg, M. Dourson. 1994. Illustration of a Threshold Model for a Combination of Three Chemicals. Conference on Chemical Mixtures and Quantitative Risk Assessment. Raleigh, NC. Sponsored by the Health Effects Research Laboratory, U.S. EPA, Raleigh, NC.

Sey, Y.M., C. Gennings, W.H. Carter, Jr., A. McDonald, L.K. Teuschler, J.E. Simmons. 1997. Nephrotoxicity of Trihalomethanes (THMs) in Female Mice. Society of Toxicology 36th Annual Meeting. March.

Simmons, J.E., C. Gennings, A. McDonald, P. Schwartz, W.H. Carter, Jr., L.K. Teuschler, Y.M. Sey. 1998. The Hepatotoxic Interaction of Chloroform ($CHCL_3$), Bromodichloromethane (BDCM), Bromoform ($CHBr_3$) and Chlorodibromomethane (CDBM). Society of Toxicology 37th Annual Meeting. March.

Simmons, J.E., C. Gennings, L.K. Teuschler, Y.M. Sey, A. McDonald. 2000. Evidence for Dose Additivity of Chloroform ($CHCl_3$) and Bromodichloromethane (BDCM) in Mice. Society of Toxicology 39th Annual Meeting, Philadelphia, PA. March.

Teuschler, L.K., C. Gennings, W.R. Hartley, H. Carter, A. Thiyagarajah, R.S. Schoeny, C. Cubbison. 1996. A Statistical Method for Detecting Departures from Additivity for Use in Estimating Toxic Interaction Effects Associated with Exposures to Toluene and Benzene Mixtures. Conference on Toxicology, Dayton OH.

Teuschler, L.K., J.E. Simmons, W.R. Hartley, A. Thiyagarajah, J.C. Lipscomb. 1999. Categorical Regression Analysis of Bromodichloromethane (BDCM) Liver Toxicity and Pathology Data. Society of Toxicology 38th Annual Meeting. March.

Thiyagarajah, A., L.K. Teuschler, JC. Lipscomb, C. Gennings, W.R. Hartley. 2000. Chronic Pathological Effects from exposure of Japanese Medaka (Oryzias Latipes) Embryos to Bromoform. Society of Toxicology 39th Annual Meeting, Philadelphia, PA. March.

INVITED TALKS

Gennings, C. 1995. Efficient Experimental Designs For Detecting And Characterizing Departure From Additivity In Mixtures Of Many Chemicals. European Conference on Combination Toxicology. Veldhoven, The Netherlands. October.

Gennings, C. 1997. Analysis and Design Issues for Combination Data using Threshold Models. Spring Meetings of the Eastern North American Region (ENAR) of the International Biometrics Society. Memphis, TN. March.

Gennings, C. 1999. Definitions of Additivity and Statistical Issues in Mixtures Analyses. Wuerzburg Forum for Toxicology. Wuerzburg, Germany. April.

Gennings, C., L.K. Teuschler, W.R. Hartley, A. Thiyagarajah, J.E. Simmons. 2000. Analysis of Ordinal Histopathology Scores in an Additivity Model. Eastern North American Region of the International Biometrics Society. March.

Hartley, W.R., C. Gennings, L. Teuschler, A. Thiyagarajah, J.E. Simmons. 1999. Advances in the Toxicological Assessment of Disinfection By-products in Rodent and Fish Biomedical Models. AWWA 1999 Annual Conference. Chicago, IL. June.

Simmons, J.E., L.K. Teuschler. 1999. Toxicology of DBPs: Water as a Complex Mixture. International Life Sciences Institute Second International Conference. The Safety of Water Disinfection: Balancing Chemical and Microbial Risks. Miami, FL. November.

Stiteler, W.M. 1999. Modeling the Joint Action of Complex Mixtures of Disinfection By-Products (DBP) in Drinking Water. Ninth Lukacs Symposium, Frontiers of Environmental and Ecological Statistics for the 21st Century: Synergistic Challenges, Opportunities and Directions for Statistics, Ecology, Environment and Society, Bowling Green, OH. April.

Stiteler, W.M. 1999. Response Addition, Dose Addition, Proportional Response Addition Methods. Peer Involvement Workshop: The Risk Assessment of Mixtures of Disinfection By-Products (DBPs) for Drinking Water Treatment Systems. Cincinnati, OH. April.

Teuschler, L.K. 1999. Research on the Risk Assessment of Mixtures of Disinfection By-Products (DBPs) in Drinking Water. Health Effects Stakeholder Meeting, Washington, DC. February.

Teuschler, L.K., J.E. Simmons. 1998. Research on the Risk Assessment of Mixtures of Disinfection By-Products (DBPs) in Drinking Water ORD/OW. Scientist to Scientist Planning Meeting, RTP, NC. October.

Teuschler, L.K., J.E. Simmons, C. Gennings, W.R. Hartley, A. Thiyagarajah. 1999. Evaluating Risks from Complex Mixtures of Disinfection By-Products (DBPs). University of Florida Symposium on Drinking Water and Health. Sarasota, FL. February.

Teuschler, L.K., J.E. Simmons, C. Gennings, W.R. Hartley, W.M. Stiteler, J.T. Colman, R.C. Hertzberg, A. Thiyagarajah, J.C. Lipscomb. 1999. A Multiple Design Approach to the Evaluation of Risks from Complex Mixtures of Disinfection By-Products (DBPs). Conference on Topics in Toxicology and Risk Assessment, Dayton, OH. April.

Teuschler, L.K., C. Gennings, W. Hartley, A. Thiyagarajah, J.E. Simmons. 1999. Novel Statistical Methods for Risk Assessment of Water Disinfection By-products. AWWA 1999 Annual Conference, Chicago, IL. June.

APPENDIX B

DETECTION OF DEPARTURES FROM ADDITIVITY IN MIXTURES OF MANY CHEMICALS WITH A THRESHOLD MODEL, *JOURNAL OF AGRICULTURAL, BIOLOGICAL AND ENVIRONMENTAL STATISTICS*, 2:198-211, 1997. (ERRATUM: *JOURNAL OF AGRICULTURAL, BIOLOGICAL AND ENVIRONMENTAL STATISTICS*, 5(2), 2000).

C. Gennings, P. Schwartz, W.H. Carter, Jr. and J.E. Simmons

Errata

Detection of Departures from Additivity in Mixtures of Many Chemicals With a Threshold Model

Chris Gennings, Pam Schwartz, W. Hans Carter, Jr., and Jane Ellen Simmons

We have recently discovered an error in the dosing values of one of the chemicals used in the mixture studied in Gennings, Schwartz, Carter, and Simmons (1997). This note provides corrected tables and figures for that manuscript. Table 1 provides a listing of dose levels (mMoles/kg), mean SDH responses, standard deviations of the responses, and sample sizes. The correction is that the dose levels of CDBM in this Table 1 are twice what they were reported in the original paper.

With the corrected dose levels, the additivity model (Equation 1.1, Gennings et al., 1997) is estimated to have a lower background parameter (β_0), which also affects the slope parameters for the other three chemicals. The quasi-likelihood ratio test of no dose-response for any of the chemicals was rejected ($p<0.001$). The threshold parameter, δ, is not significantly different from zero (Table 2, $p=0.953$). The point estimates and 95% confidence intervals for the threshold for each chemical are provided in Table 3. All four intervals include zero. The fitted-curves from the additivity model for each single chemical are provided in a new Figure 3.

Using the additivity model given in Equation (1.1) and parameter estimates in Table 2, the predicted SDH response under additivity at $\mathbf{x}=(0.208, 0.084, 0.568, 0.012)$ for (BDCM, CDBM, $CHCl_3$, $CHBr_3$) (i.e., the Krasner mixture) is $\hat{y} = 40.5$ with standard deviation of 2.46. The large sample 95% interval constructed under the assumption of additivity associated with $\mathbf{x}$ is [31.9, 51.9]. Here, the observed sample mean response, $\overline{y} = 43.9$, is included in the prediction interval. Therefore, these data provide no evidence of departure from additivity at the combination point of interest.

References

Gennings, C., Schwartz, P., Carter, Jr., W. H., Simmons, J. E. (1997). "Detection of Departures From Additivity in Mixtures of Many Chemicals with a Threshold Model," *Journal of Agricultural, Biological, and Environmental Statistics* 2, 198-211.

Table 1: Observed Dose-Response Data

Chemical	Dose (mMoles/kg)	Mean SDH	Standard Deviation	Sample Size
BDCM				
	0	19.3	3.77	10
	0.152	21.4	3.44	10
	0.305	26.6	7.97	11
	0.76	39.6	12.4	11
	1.52	154.4	133.8	10
	3.05	187.4	163.9	8
CDBM				
	0	21.2	3.07	13
	0.304	25.6	5.09	16
	0.610	32.0	9.76	18
	1.52	167.9	112.1	14
	3.04	146.7	123.9	5
	6.10*	-	-	-
$CHCl_3$				
	0	15.6	3.27	12
	0.152	16.8	4.20	13
	0.305	21.3	5.85	15
	0.76	30.3	19.9	15
	1.52	50.8	22.5	13
	3.05	80.2	9.35	12
$CH\ Br_3$				
	0	21.3	5.97	16
	0.152	24.1	6.77	14
	0.305	23.8	4.14	15
	0.76	38.2	25.1	16
	1.52	87.6	21.8	10
	3.05	130.0	125.0	14

*All mice that received 6.10 mMoles/kg of CDBM died before the end of the 14 day dosing period.

Table 2: Estimated Model Parameters for the Additivity Model Associated with the Power Parameter of l=0.5

Parameter*	Estimate	p value	Threshold Estimates
b_0	4.42	<0.001	
b_1 (BDCM)	3.57	<0.001	0.007
b_2 (CDBM)	3.84	<0.001	0.007
b_3 ($CHCl_3$)	1.54	<0.001	0.017
b_4 ($CHBr_3$)	2.44	<0.001	0.011
d	0.0256	0.953	

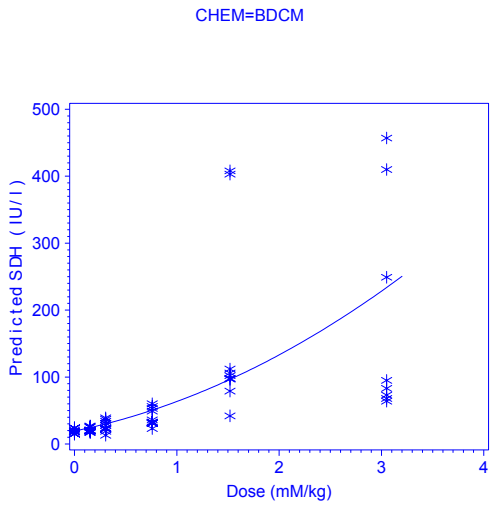
* The estimate for τ is 28.1.

Table 3: Point and Interval Estimates for the Threshold for Each Chemical

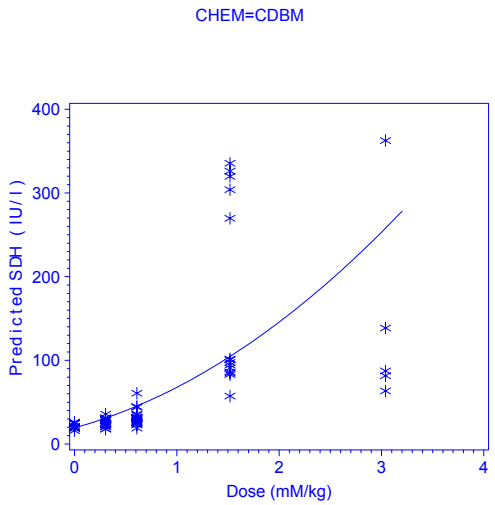
Chemical	Threshold Estimate	95% Confidence Interval
BDCM	0.007	[0, 0.248]
CDBM	0.007	[0, 0.231]
$CHCl_3$	0.017	[0, 0.573]
$CHBr_3$	0.011	[0, 0.362]

Figure 3: Observed Data (*) and Predicted Responses for the Additivity Surface for Each Single Chemical: (A) BDCM, (B) CDBM, (C) Chloroform, (D)Bbromoform. The fitted regression is based on the square root-linear model given in (1.1), with parameter estimates provided in Table 2.

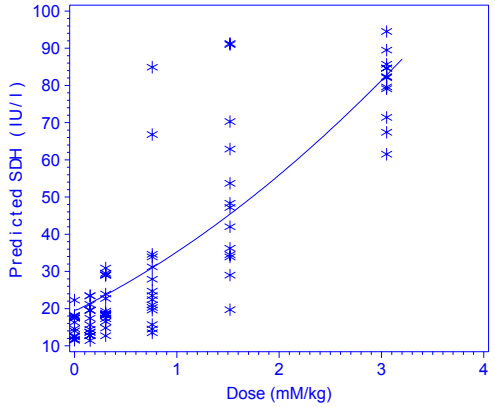
A



B



C



D