



Air Force Research Laboratory

The Toxicology of Chemical Mixtures Risk Assessment for Human and Ecological Receptors

Teresa R. Sterner

**Operational Technologies, Inc.
1370 N. Fairfield Road, Suite A
Beavercreek OH 45432**

Peter J. Robinson

**Alion Science and Technology
P O Box 31009
Dayton OH 45437-0009**

David R. Mattie

**Air Force Research Laboratory
Human Effectiveness Directorate
Applied Biotechnology Branch
Wright-Patterson AFB, OH 45433-5707**

G. Allen Burton

**Institute for Environmental Quality
Wright State University
3640 Colonel Glenn Hwy
Dayton OH 45435**

October 2005

FINAL REPORT FOR THE PERIOD JANUARY 2004 TO OCTOBER 2005

Approved for public release;
distribution unlimited

20061212226

**Air Force Research Laboratory
Human Effectiveness Directorate
Biosciences and Protection Division
Applied Biotechnology Branch
Wright-Patterson AFB, OH 45433-5707**

NOTICES

When US Government drawings, specifications or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from the Air Force Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with the Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Service
8725 John J. Kingman Rd., Ste 0944
Ft Belvoir, Virginia 22060-6218

DISCLAIMER

This Technical Report is published as received and has not been edited by the Technical Editing Staff of the Air Force Research Laboratory.

TECHNICAL REVIEW AND APPROVAL

AFRL-HE-WP-TR-2005-0173

The animal use described in this study was conducted in accordance with the principles stated in the "Guide for the Care and Use of Laboratory Animals", National Research Council, 1996, and the Animal Welfare Act of 1966, as amended.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE DIRECTOR

//SIGNED//

MARK M. HOFFMAN

Deputy Chief, Biosciences and Protection Division
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) October 2005		2. REPORT TYPE Final Technical Report		3. DATES COVERED (From - To) January 2004 - October 2005	
4. TITLE AND SUBTITLE The Toxicology of Chemical Mixtures Risk Assessment for Human and Ecological Receptors				5a. CONTRACT NUMBER F33615-02-P-A271	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 62202F	
				5d. PROJECT NUMBER 2312A	
5. AUTHOR(S) *Teresa R. Sterner, **Peter J. Robinson, ***David R. Mattie, ****G. Allen Burton				5e. TASK NUMBER 2312A2	
				5f. WORK UNIT NUMBER 2312A215	
				6. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Operational Technologies Dayton OH *Alion Science and Technology Corp P.O. Box 31009, Dayton OH 45437 ***AFRL/HEPB, WPAFB OH ****Institute for Environmental Quality Wright St Univ, Dayton OH	
7. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Materiel Command*** Air Force Research Laboratory Human Effectiveness Directorate Biosciences and Protection Division Wright-Patterson AFB OH 45433				8. PERFORMING ORGANIZATION REPORT NUMBER	
				10. SPONSOR/MONITOR'S ACRONYM(S) HEPB	
2. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited. Cleared by AFRL/WS-06-0570, 28 Feb 06.				11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-HE-WP-TR-2005-0173	
				3. SUPPLEMENTARY NOTES	
4. ABSTRACT Mixtures of chemicals represent a concern in all areas of toxicology and a challenge to risk assessors as most sites with environmental contamination, including Department of Defense sites, involve simultaneous or sequential exposures of multiple chemicals to human or ecological receptors. A major concern with mixtures is that they may lead to outcomes (health effects) or increased toxicity (synergism) not expected from the risk characterization of individual chemicals at the site. The focus of this literature review is to examine the principles used to determine toxicity in chemical mixtures risk assessment for human and ecological receptors at hazardous waste sites. This review provides an overview of current practices useful to a remedial project manager or risk assessor prior to their first chemical mixtures risk assessment.					
5. SUBJECT TERMS Military Logistics, Complex Systems, Asymmetric Command and Control, Decision-making, Collaborative Sensemaking, Information Management, Automation					
6. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES 102	19a. NAME OF RESPONSIBLE PERSON Dr. David R. Mattie
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED			19b. TELEPHONE NUMBER (include area code) 937-904-9569

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

PREFACE

This review was accomplished at the Applied Biotechnology Branch, Human Effectiveness Directorate of the Air Force Research Laboratory, Wright-Patterson AFB, OH, under John J. Schlager, Branch Chief. The contractors were under the direction of Peter Lurker (Operational Technologies Corp., Dayton, OH, Contract # F33601-02-F-A211) and Darol Dodd (Alion Science and Technology, Inc. (formerly ManTech Environmental Technology, Inc.), Dayton, OH, Contract # F-33615-00-C-6060). This technical report was written for AFRL Workunit 2312A215.

This work was performed to support development of a JP-8 mixtures model and to fulfill obligations toward a Masters Degree in Environmental Science at Wright State University, Dayton, OH (Teresa R. Sterner).

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

INTRODUCTION	1
BACKGROUND	2
Mixtures Risk Assessment Requirements	2
Mixtures Complexity and Issues	3
Mechanisms of Interaction	6
Review Perspective on HHRA and ERA	7
TOXICOLOGY OF CHEMICAL MIXTURES RISK ASSESSMENT FOR HUMAN	
RECEPTORS	9
Incorporation of Mixtures in HHRA.....	9
Risk Calculation Procedures for Whole Mixture Approaches	11
Mixture of Concern Procedure.....	12
Sufficiently Similar Mixture Procedure	14
Comparative Potency Procedure.....	15
Risk Calculation Procedures for Mixture Component Approaches	18
Response Addition Approach.....	20
Dose Addition Approaches	23
<i>Hazard Index Procedure.....</i>	<i>24</i>
<i>Fraction Surrogate Procedures.....</i>	<i>28</i>
<i>Relative Potency Factor/Toxic Equivalency Factor Procedure</i>	<i>33</i>
Interaction Approaches	37
Mixture Weight of Evidence Procedure	38
Interaction Hazard Index Procedure.....	44
Pharmacokinetic/Pharmacodynamic Modeling	48
TOXICOLOGY OF CHEMICAL MIXTURES RISK ASSESSMENT FOR ECOLOGICAL	
RECEPTORS	55
Incorporation of Mixtures in ERA	55
Risk Calculation Procedures for Component Dose Addition Approaches.....	59
Hazard Index Procedure	63
Toxic Equivalent Factor Procedure	69
Risk Procedures for Whole Mixture Approaches.....	71
Apparent Effects Threshold Procedure.....	77
CONCLUSIONS	79
REFERENCES	83

LIST OF TABLES AND FIGURES

Table 1. Advantages and Disadvantages of the Mixture of Concern Procedure	14
Table 2. Advantages and Disadvantages of the Sufficiently Similar Mixture Procedure	15
Table 3. Advantages and Disadvantages of the Comparative Potency Procedure	17
Table 4. Advantages and Disadvantages of the Response Addition Approach	23
Table 5. Advantages and Disadvantages of the Hazard Index Procedure	28
Table 6. Advantages and Disadvantages of the Fraction Surrogate Procedures	32
Table 7. Advantages and Disadvantages of the TEF Procedure	37
Table 8. Mixture Weight of Evidence Alphanumeric Classification and Binary Weight of Evidence Scores	40
Table 9. Advantages and Disadvantages of the Mixture WOE Procedure	44
Table 10. Interaction HI Weight of Evidence Scores (B)*	46
Table 11. Advantages and Disadvantages of the Interaction HI Procedure	48
Table 12. Advantages and Disadvantages of PBPK/PBPD Modeling	54
Table 13. Comparison of U.S. EPA Steps for ERA and DoD Tiered ERA	56
Table 14. Advantages and Disadvantages of the Hazard Index Procedure	68
Table 15. Advantages and Disadvantages of the TEF Procedure	71
Table 16. Advantages and Disadvantages of Whole Mixture Laboratory and Field Toxicity Testing Procedures	77
Table 17. Advantages and Disadvantages of the AET Procedure	79
 Figure 1. Example Isobole for Two Hypothetical Chemicals	 20

LIST OF ACRONYMS

2,3,7,8-TCDD	2,3,7,8-tetrachlorodibenzo-[p]-dioxin
ACGIH	American Council of Government Industrial Hygienists
AET	apparent effects threshold
Ah	aryl hydrocarbon
ATSDR	Agency for Toxic Substances and Disease Registry
B	weight of evidence score or factor
BINWOE	binary weight of evidence
BMD	benchmark dose
BTEX	benzene, toluene, ethylbenzene and xylenes
CYP1a	cytochrome P-450-1a
DoD	Department of Defense
DoE	Department of Energy
EC ₅₀	effective concentration for 50% of the exposed organisms
ED ₁₀	10% effective dose
ERA	ecological risk assessments
EROD	ethoxyresorufin-O-deethylase
f	exposure factor
GI	gastrointestinal
HHRA	human health risk assessments
HI	hazard index
HQ	hazard quotient
IEF	induction equivalency factor
IRIS	Integrated Risk Information System
LOAEL	lowest observed adverse effect level
LOEC	lowest observable effect concentration
M	magnitude factor
MA DEP	Massachusetts Department of Environmental Protection
NOAEL	no observed adverse effect level
NOEC	no observed effect concentration
NPL	National Priority List
OSHA	Occupational Safety and Health Administration
PAH	polycyclic aromatic hydrocarbon
PBPD	pharmacodynamic
PBPK	physiologically based pharmacokinetic
PCB	polychlorinated biphenyl
PD	pharmacodynamic
PEL	permissible exposure level
PK	pharmacokinetic
PNEC	predicted no effect concentrations
PPRTV	Provisional Peer Reviewed Toxicity Values
QSAR	quantitative structure activity relationship
RAGS A	Risk Assessment Guidance for Superfund part A
RCRA	Resource Conservation and Recovery Act
RfC	reference concentration
RfD	reference dose
RPF	relative potency factor
TCDD	2,3,7,8-tetrachlorodibenzo-[p]-dioxin
TCE	trichloroethylene

LIST OF ACRONYMS (CONTINUED)

TE	toxic equivalent
TEEL	temporary emergency exposure limit
TEF	toxic equivalency factor
TEQ	toxic equivalent
TIE	toxicity identification evaluation
TLV	threshold limit value
TLV	threshold limit value
TPH	total petroleum hydrocarbon
TPHCWG	Total Petroleum Hydrocarbon Criteria Working Group
TRV	toxicity reference value
TTC	target organ toxicity concentration
TTD	target organ toxicity dose
U.S. EPA	U.S. Environmental Protection Agency
UF	uncertainty factor
USACHPPM	U.S. Army Center for Health Promotion and Preventative Medicine
WET	whole effluent toxicity
WOE	weight of evidence

THE TOXICOLOGY OF CHEMICAL MIXTURES RISK ASSESSMENT FOR HUMAN AND ECOLOGICAL RECEPTORS

INTRODUCTION

Mixtures of chemicals represent a concern in all areas of toxicology, from drug and herbal supplements interaction research (Abebe, 2002) to infants' and children's' developmental and neurological health (Tilson, 2000) or effluent testing under the Clean Water Act (U.S. EPA, 2002a). Mixtures toxicology is a challenge to risk assessors as most sites with environmental contamination involve simultaneous or sequential exposures of multiple chemicals to the receptors, human or ecological (U.S. EPA, 1986). Hazardous waste sites under U.S. Environmental Protection Agency (U.S. EPA) supervision can contain 100 or more identified chemicals of concern (Johnson and DeRosa, 1995). A major concern with mixtures is that they may lead to outcomes (health effects) or increased toxicity (synergism) not expected from the risk characterization of individual chemicals at the site (Connolly, 2001).

Chemical mixtures toxicity has been and will continue to be a concern to the Department of Defense (DoD). Past and present contaminated sites number more than 28,000 and are located on over 11,000 DoD installations and properties, including 155 sites proposed or included on the National Priority List (NPL). The majority of these sites have been closed or restored; fewer than 30% remain in some state of assessment or remediation (Wireman *et al.*, 2003). Many of these sites will be impacted by a combination of chlorinated solvents (e.g., trichloroethylene (TCE) and perchloroethylene) and petroleum products (e.g., jet fuels, diesel, gasoline, aviation gasoline), with additional chemicals likely to be found as well (Lebron, 2003). The Agency for Toxic Substances and Disease Registry (ATSDR) is also focusing on frequently occurring mixtures at NPL sites and has included mixtures toxicity as part of their study agenda through the year 2010 (Spengler and Falk, 2002).

The focus of this literature review is to examine the principles used to determine toxicity in chemical mixtures risk assessment for human and ecological receptors at hazardous waste sites. The review provides an overview of current practices useful to a remedial project manager or risk assessor prior to their first chemical mixtures risk assessment.

BACKGROUND

Mixtures Risk Assessment Requirements

The definition of mixtures risk assessment is necessarily vague: "an analysis, characterization and possible quantification of the combined risks to health or the environment from multiple agents or stressors". Likewise, the prescription for when a mixtures risk assessment should be conducted is also unclear: "whenever the combined impact of multiple stressors needs to be considered" (U.S. EPA, 2002b). Mixtures investigations for human health have been advised for many years; the need to consider mixture toxicity was stated in the original Risk Assessment Guidance for Superfund part A (RAGS A) (U.S. EPA, 1989), the Clean Air Act in 1990, the Safe Drinking Water Act of 1996 (Yang, 1998) and the Food Quality Protection Act of 1996 governing pesticide safety (U.S. EPA, 2002c). The toxicity of manufactured and generated complex mixtures such as coke oven emissions and jet fuels JP-5 and JP-8 have been evaluated by the U.S. EPA (2004a) and ATSDR (1998), respectively. Environmental impacts of cumulative stressors has been recognized for years; cumulative effects analysis was written into the National Environmental Policy Act of 1969 (U.S. EPA, 2002b). Ecological risk assessments on DoD sites are expected to consider multiple stressors, including chemical, physical and biological (Wentzel *et al.*, 1994; 1996). From the regulations and guidance above, some level of mixtures risk assessments is indicated for nearly every complex site.

However, not every mixtures risk assessment will be the same. From the definition, it is apparent that not all assessments will be quantitative (U.S. EPA, 2002b). A screening level assessment may be sufficient to determine that it is not worth investing resources in a more data intensive assessment (ILSI, 1999). Such an assessment may indicate that one chemical drives the cumulative risk and that chemical has already been determined to pose unacceptable risk to the receptors (U.S. EPA, 2002c). Furthermore, the type of mixtures assessment performed will be somewhat determined by the needs of the risk manager and limitations posed by the available data (ILSI, 1999). For the majority of environmental mixtures, there are insufficient data available to be able to accurately estimate risk of component interactions (Teuschler and Hertzberg, 1995). However, sufficient mixtures exposure and toxicity data can help to reduce uncertainty and increase confidence in the risk assessment (ILSI, 1999).

Chemical mixtures risk assessments need not include every contaminant past or present at the site. Exposure pathways need to be complete (i.e., the chemical reaches the receptor). If the chemical is no longer present in the environment, if the chemical will not come into contact with a receptor or if the chemical is not taken up by a particular route (e.g., non-volatile chemicals are not inhaled), then the chemical should not be included in the assessment (U.S. EPA, 2002c; Mumtaz *et al.*, 1998; Simini *et al.*, 2000). Although all contaminants and routes of exposure should be assessed qualitatively, scarce resources should be focused on chemicals and exposure scenarios identified to have potentially higher risks to the receptors for quantitative assessment (U.S. EPA, 2002c).

Mixtures Complexity and Issues

"Complexity is a major reason why mixtures have not been well studied" (Carpenter *et al.*, 2002). Some mixtures are simple, consisting of a relatively small number of components (usually considered as ten or fewer). Simple mixture components are readily identifiable and usually quantifiable. However, most environmental mixtures are considered complex mixtures,

meaning they include tens or up to thousands of individual components. The composition of complex mixtures usually is not fully known or quantifiable (Feron and Groten, 2002; Groten, 2000). Some environmental contaminants start out as manufactured complex mixtures of congeners (e.g., polychlorinated biphenyls (PCBs), petroleum products) or generated compounds (e.g., drinking water disinfection by-products) while other mixtures are coincidental, released from different sources but transported through the same media, resulting in a simultaneous or sequential exposure (ATSDR, 2001a).

Definitions of toxicological interactions between components can also be complex. Differing definitions of synergism, antagonism and "no interaction" in published literature can be problematic when trying to assess a chemical's potential for interaction. Synergism and antagonism are comparative terms dependent on the definition of "no interaction" (Hertzberg and MacDonell, 2002). The U.S. EPA's definition of "no interaction" is either dose addition (additivity) or response addition (independence) (U.S. EPA, 2000), both of which will be addressed later in this review. The type of "no interaction" expected must be clarified in a study as dose and response addition do not give the same results (NORA, 2004; Simmons, 1995). A synergistic effect is a response statistically greater than expected from a linearized dose-response curve (Lang, 1995) or a response unexpected from the study of the individual mixture components (Mauderly, 1993). An antagonistic effect is lower in toxicity than expected from a dose-response relationship; however, lower toxicity is not to be confused with non-toxic (Hertzberg and MacDonell, 2002; Lang, 1995). Because additivity means "no interaction" in most studies but synergism in a few and because there are multiple synonyms for the other terms as well, statistical procedures need to be well explained in published studies and the magnitude of interaction between chemicals reported (Hertzberg and MacDonell, 2002). Additional barriers to interpreting studies can stem from inappropriate study design and statistical analysis (Deneer, 2000; U.S. EPA, 2000).

Chemical characteristics add additional complexity. Individual compounds generally act upon one or more cellular receptors in multiple tissues (Carpenter *et al.*, 2002). A single component can have age dependent effects; lead has little effect on adults at concentrations that are irreversibly neurotoxic to children (Carpenter *et al.*, 1998). Toxic agents may have different effects depending on the route of exposure; asbestos fibers are carcinogenic only when inhaled (Carpenter *et al.*, 2002). Toxicity of chemicals (e.g., explosives) may change with environmental conditions, including pH, organic matter in the soil and composition of the soil (Wentzel *et al.*, 1994; 1996). Compounds may be metabolized, effectively becoming a mixture of parent, intermediates and metabolites (Carpenter *et al.*, 1998; Yang *et al.*, 1998). Mixtures of chemicals may have synergistic responses in one target organ, but additive or antagonistic responses in a different organ (Carpy *et al.*, 2000). These factors combined make mixture toxicity more difficult to study, but not impossible (Yang, 1998).

Exposure considerations are not simple with mixtures either. Receptors are exposed to mixtures continually, not only from the environment, but also through food and water. Humans are exposed through home, work, hobbies, pharmaceuticals and smoking (Viau, 2002). Receptors are exposed to mixtures not just simultaneously, but also in succession. These exposures are considered mixtures exposures when persistent chemicals are retained and concentrated in the body (Health Council of the Netherlands, 2002). A simultaneous exposure can also occur when a metabolite from a previous exposure is present in the body during the next exposure event. Sequential exposures increase the potential for mixtures interactions (Mauderly, 1993).

Mixtures change over time and distance at contaminated sites (Schulte, 2003). Manufactured mixtures such as petroleum products vary in composition from different sources (crude oils) and over time from the same source (Seed *et al.*, 1995) prior to being introduced into the environment. Mixture components may partition into different environmental compartments (surface water, sediment, pore water, groundwater), which are transported at

different rates. Compounds undergo transformations differentially (U.S. EPA, 1986). Potential transformations include photolysis, hydrolysis, degradation and biodegradation. For example, PCBs are subject to partitioning, some biodegradation and, for some congeners, bioaccumulation (Teuschler *et al.*, 2001). Bioavailability of chemicals changes with these physical processes. Polycyclic aromatic hydrocarbons (PAHs) such as naphthalene are persistent but become less bioavailable over time (155 days) to earthworms due to sequestration in the soil (Kelsey and Alexander, 1997). Following an environmental release, receptors will be exposed to a subset of the original mixture with the addition of transformation by-products. Health criteria based on the original mixture may not be representative of toxicological effects from the weathered mixture (ATSDR, 2001a).

Mechanisms of Interaction

Multiple interactions in the environment can affect the well-being of receptors. These can include chemical-chemical, chemical-biological, chemical-physical, physical-biological, etc. Although chemical-physical and other interactions can change toxicity (e.g., ultraviolet light increases toxicity of PAHs, heat accelerates breathing rate and uptake of volatiles, toluene increases noise induced hearing loss) (Foran and Ferenc, 1997; NORA, 2004; U.S. EPA, 2002b). This review will focus on chemical-chemical interactions.

Chemicals can interact to change toxicity in three fundamental ways. The first is simple chemical reaction. Two chemicals can react in the environment to form a third compound having a different toxicity than the parent compounds. Reaction can increase or decrease toxicity (ATSDR, 2001a; U.S. EPA, 1986); chelating chemicals complex with metal ions, forming insoluble complexes with much decreased toxicity (Calabrese, 1995).

The second and third types of interactions are more complex. Pharmacokinetic (PK) interactions occur during the processes of absorption, distribution, metabolism or elimination (U.S. EPA, 2000). The result of PK interactions is an increase or decrease in the amount of

chemical in the tissue where the toxicological effect takes place (Lang, 1995). Pharmacokinetic interactions may result in antagonistic or synergistic changes in effects compared to effects expected from single component toxicity (Wentsel *et al.*, 1994; 1996). Tissue absorption can be decreased through competition for cellular receptors or increased by induction of transport pathways (U.S. EPA, 2000). Gastrointestinal (GI) absorption can be altered by disruptions in the acid-base balance of the stomach and intestines or by changes in the motility of the GI tract (Calabrese, 1995). Lipophilic compounds may increase absorption of hydrophilic compounds normally absorbed at lower rates (Zeliger, 2003). Distribution can be altered by plasma protein binding (Calabrese, 1995); metals distribution can be changed due to metallothionein binding. Metabolism can be altered through induction or depletion of metabolic enzymes, such as cytochrome P450 induction or glutathione depletion (U.S. EPA, 2000). Elimination of weak acids and bases is affected by kidney pH (Calabrese, 1995). Due to the number of ways PK interactions can take place, it is not surprising that most chemical interactions are found to be pharmacokinetic (Hertzberg and MacDonell, 2002; Robinson and MacDonell, 2004).

Pharmacodynamic (PD) or toxicodynamic interactions take place at the site of toxic action (U.S. EPA, 2000). The tissue response to a delivered dose of the first chemical is changed due to the presence of the second or additional chemicals (Lang, 1995). Examples of PD interactions can include depletion of protective factors within cells, increases or decreases of tissue repair rates and changes in immune responses (U.S. EPA, 2000).

Review Perspective on HHRA and ERA

Although human health risk assessments (HHRA) and ecological risk assessments (ERA) have many of the same data needs and could perhaps be handled more efficiently if these data were shared (Suter, 2004), the objectives of human and ecological assessments differ. HHRA focus on the protection of individuals, demanding acceptable levels of potential risk to individuals from the most sensitive subpopulation. Carcinogenicity is a major concern.

ERAs are generally population defined assessments. Mixtures are evaluated for their impact on the total population (Suter, 2004; U.S. EPA, 2002b) with endpoints of survivability, growth and reproduction. Carcinogenicity is rarely evaluated. Only threatened and endangered species are evaluated on an individual level (Wireman, 2004, personal communication). Due to these different objectives and the risk calculations used to meet these objectives, the toxicology of mixtures risk assessment is considered separately for human and ecological receptors in this review.

Another difference between human and ecological risk assessment of mixtures lies in the term, "weight of evidence". In both risk assessment realms, weight of evidence (WOE) refers to an overall concept for evaluating site risk and uncertainty, using all pertinent information (U.S. EPA, 2004b). A lone set of experiments, which in turn lead to a conclusion regarding a single endpoint, is not enough evidence to characterize risk for an entire contaminated site (Fairbrother, 2003). Instead, the weight of evidence process combines the outcomes of several lines of evidence to make conclusions about the health risks to humans or the ecosystem (Burton *et al.*, 2002). WOE decisions, based on professional judgment, take into consideration the quality, adequacy and consistency of the data used in the site or chemical-specific risk assessment (U.S. EPA, 2004b). However, in HHRA, a procedure also called "weight of evidence" (discussed in the Interaction Approaches section) was created to characterize potential chemical interaction. To clarify the difference between the general risk assessment practice and the specific interaction procedure, the term "mixture WOE procedure" is used throughout this review for the HHRA tool.

TOXICOLOGY OF CHEMICAL MIXTURES RISK ASSESSMENT FOR HUMAN RECEPTORS

Incorporation of Mixtures in HHRA

The U.S. EPA Framework for Cumulative Risk Assessment (2002b) projects three main phases for cumulative risk assessments: planning/problem formulation, analysis and risk characterization. These phases mesh with the original risk assessment paradigm (U.S. EPA, 1989). The activities of the problem formulation phase, which includes defining the purpose and scope of the assessment, incorporate the hazard identification portion of the paradigm. The analysis phase includes the dose-response assessment and exposure assessment portions of the paradigm. The risk characterization phase equates to the risk characterization portion of the paradigm. The 2002 Framework does not change the risk assessment; it suggests additional considerations for incorporating mixtures risk. For example, during the analysis phase, risk assessors are to consider prior exposures, susceptibilities or differences in the abilities to recover from exposure that may make a sub-population more vulnerable to mixtures toxicity (U.S. EPA, 2002b). Mixtures toxicology is involved in both the dose-response assessment (analysis phase) and risk characterization portions of the paradigm (ILSI, 1999).

A point stressed in the risk characterization phase of the 1986 Chemical Mixture Guidelines and the 2002 Framework is the necessity of dealing with uncertainty. After describing the quantitative (if possible) or qualitative results of the mixtures HHRA, major assumptions, limitations, possible bias and uncertainties of these results are to be discussed thoroughly (U.S. EPA, 1986; 2000; 2002b). This requirement has been described as a "major communication challenge" (U.S. EPA, 2002b). General assumptions and uncertainties associated with the methods of risk calculation are listed in this review.

The risk calculation procedure selected for use in an HHRA is generally dictated by the availability of quality data for the mixture at the site (Teuschler *et al.*, 2001; U.S. EPA, 1986). The use of several approaches is recommended; the results of approaches followed to

conclusion can be evaluated as a range of risk estimates for the mixtures toxicity of the site. Should the exposure or toxicological data quality be poor, a qualitative assessment may be performed, describing all available information on the mixture, its components and potential interactions (U.S. EPA, 2000).

For mixtures risk assessment, data requirements are comparable with individual chemical assessments. Human epidemiological or clinical studies are preferred, especially for whole mixture toxicity. The best possible situation would be a human study using the mixture of concern, where the health effects seen were causally attributed to the mixture and the exposure occurred at site relevant concentrations and through a route linked to a complete exposure pathway. A comprehensive animal database can be used if it is supported by animal studies on the mixtures or extracts or by human and animal data on the most toxic and most prevalent components of the mixture. Toxicity studies using environmentally relevant concentrations and routes of exposure are of more value (U.S. EPA, 2000). Studies and assays determining mode of action, a general description of how the toxicant causes effects (Andersen and Dennison, 2004), are valuable for evaluating mixtures risk. Preferred exposure data include information on the stability of the mixture in the environment, the components remaining from the mixture in every completed exposure pathway and the bioavailability of mixture components at the site (U.S. EPA, 2000).

Procedures for assessing toxicity in mixtures HHRA can be split into two types. Whole mixture approaches are considered top-down methods used for existing (manufactured and generated) mixtures. The component approach, or bottom-up method, is useful at hazardous waste sites or in occupational settings, where the mixture is incidental (Borgert *et al.*, 2003; Groten, 2000). Should the choice of method be unclear, multiple methods may be used and the range of risk estimates compared (Teuschler *et al.*, 2001). A combination of whole mixture and component methods may even strengthen the overall mixtures HHRA (U.S. EPA, 2000).

Risk Calculation Procedures for Whole Mixture Approaches

Dose-response assessments and risk characterizations can be performed for mixtures in their entirety. The advantage is obvious; health effects, including interactions, associated with all components of the mixture can be determined. Mode of action data are not necessary (U.S. EPA, 2000) as mode of action and chemical interaction are not formally considered but are part of the whole exposure resulting in toxicological effects (U.S. EPA, 1986). Any interactions among the components should be represented by the health effects observed (ATSDR, 2001a). However, this lack of component and interaction data can compromise the health assessment; for example, if toxicity of the mixture is dominated by a single or a few components and those components are not present in the weathered mixture (e.g., volatilized), then risk values will overestimate the health effects (Krishnan *et al.*, 1997).

Disadvantages to whole mixtures assessments are equally obvious. Mixtures of the same name are often not identical; petroleum products vary by crude source and over time (Seed *et al.*, 1995) and are only required to remain within product performance specifications. Although whole mixture assessment is recommended for drinking water disinfection by-products (Teuschler and Simmons, 2003; Teuschler *et al.*, 2000), components of these mixtures vary by season and distribution system (Krishnan *et al.*, 1997). Similarly, the assessor assumes that the original mixture used in the toxicity study is analogous to the environmental mixture (ATSDR, 2001a). Mixture composition changes over time in the environment; a whole mixture assessment will not be assessing risk for the mixture present at the site. In a fairly stable mixture, relative proportions of the chemicals remain roughly constant and whole mixture data are considered valid. Professional judgment is necessary to determine the stability of the mixture, its bioavailability and the applicability of the whole product data to the weathered mixture (U.S. EPA, 2000).

Another disadvantage of whole mixtures assessments is the lack of data. Available data are frequently inadequate to evaluate sensitive endpoints, such as development and

reproduction (U.S. EPA, 2000). Use of *in vitro* assays may augment the available data. Although assay endpoints can include cytotoxicity (Malich *et al.*, 1998), mutagenicity and genotoxicity, not all toxicological endpoints can be assessed through *in vitro* methods (Donnelly *et al.*, 1995).

Mixture of Concern Procedure

The U.S. EPA (2000) preferred approach for mixtures risk characterization is to use subchronic or chronic toxicity data from the mixture of concern as if the mixture were a single compound. This is also the routine practice of the European Union (Health Council of the Netherlands, 2002). The ATSDR has employed this approach for PCBs, based on the toxicity of Aroclor 1254, and petroleum products such as jet fuels and kerosene (Pohl *et al.*, 1997). The U.S. EPA's Integrated Risk Information System (IRIS) includes a toxicological assessment for Aroclor 1016 which is considered of medium confidence, acknowledging that the congeners found in the environment do not match commercial mixtures of PCBs (U.S. EPA, 2000; 2004a). Complete mixture data are more available for complex manufactured or generated mixtures due to the large quantities produced (U.S. EPA, 2000).

Data needed for a mixture of concern dose-response toxicity assessment are the same as for any single chemical (U.S. EPA, 2000). As such, a complete database would consist of chronic human or animal toxicity data with supporting research including reproductive and developmental studies in two species (U.S. EPA, 1989). Again, human epidemiological or clinical studies are preferred. Supporting studies may be based on extracts, although extracts may be better handled as sufficiently similar mixtures (discussed below) (U.S. EPA, 2000). Extraction procedures such as reverse osmosis are capable of concentrating water without loss of volatile components (Simmons *et al.*, 2002), maintaining the relative proportionality of the mixture. The toxicity assessment should consider the differences that environmental fate and bioavailability may have on the mixture, using available data (U.S. EPA, 2000).

The main advantage of the mixture of concern approach is the accommodation of real life exposures (Table 1). Aside from the whole mixture advantages already discussed, whole mixture risk value calculation follows the same methods (reference dose (RfD), reference concentration (RfC), cancer slope factor) as are used for individual chemicals, so the calculations are relatively simple and familiar (U.S. EPA, 2000). Of course, the same limitations apply as well, including uncertainty in species to species extrapolation in animal studies and exposure concentration and duration extrapolations in epidemiological studies (U.S. EPA, 1986). Data from different routes of exposure can be used with the same stipulations (e.g., portal of entry effects are discounted) as individual chemicals. Similarly, the risk value is derived with only toxicity information; site exposure concentrations are not needed for this calculation and the value should be useful for all sites contaminated with this mixture (U.S. EPA, 2000).

The large number of whole mixtures for potential testing is a prime disadvantage of the mixture of concern method (Borgert *et al.*, 2003). Extrapolation from high to low doses can also be questionable, due to differential metabolism (Feron *et al.*, 2002). Reproducibility of results is more difficult at low mixture concentrations (Carpy *et al.*, 2000). As mentioned above, mixtures change in the environment; confidence in a risk value is decreased if the mixture is known to change following release. Data on the toxicity of additional mixtures similar to the original or to the weathered mixtures can help increase confidence in a risk value (U.S. EPA, 1986).

Table 1. Advantages and Disadvantages of the Mixture of Concern Procedure

Advantages	Disadvantages
• Incorporates components, interactions, mechanisms of action • Accommodates real life exposures • Same toxicological data requirements as for individual chemical data • Site exposure data not needed • Same risk value calculations as for individual chemicals	• Original mixture composition varies • Mixture composition changes in the environment • Lack of toxicity data • Large number of mixtures to be tested • High to low dose extrapolation

Sufficiently Similar Mixture Procedure

It is also acceptable to use toxicity information from mixtures similar to the mixture of concern. This procedure should be used when toxicity data are unavailable for the mixture of concern but are available for a closely related mixture (U.S. EPA, 2000). This approach assumes biological activity of the mixture will be similar because the chemical makeup is similar (NORA, 2004). The mixtures should have the same or nearly the same components, in similar proportions; known health effects and component fate and transport parameters should also be similar for the two mixtures (ATSDR, 2001a). Any available information should be used to compare the mixtures and their components; differences in pharmacokinetics, toxicity or bioavailability should be considered in the similarity decision (U.S. EPA, 2000). Computerized pattern-recognition techniques using gas chromatography/mass spectrophotometer fingerprinting can assist in the chemical constituent comparison of mixtures (Health Council of the Netherlands, 2002). Aside from chemical component data, sufficient toxicity data on the similar mixture must be available to support a risk value with adequate confidence to be used in risk characterization (U.S. EPA, 2000).

The advantages to using similar mixtures are the same as for mixture of concern (Table 2). The added benefit is that existing toxicological data on the similar mixture can be used without requiring further study. The disadvantage to this is the rare availability of similar mixture

data (NORA, 2004; Seed *et al.*, 1995; U.S. EPA, 2000). Due to this deficiency, the sufficiently similar mixture procedure has not been used extensively and has been primarily applied to carcinogenic effects (NORA, 2004). Furthermore, extensive professional judgment is necessary to determine really how similar the mixtures are (Seed *et al.*, 1995). The conclusion of sufficient similarity must be fully supported and uncertainties analyzed in the risk assessment document (U.S. EPA, 2000).

Table 2. Advantages and Disadvantages of the Sufficiently Similar Mixture Procedure

Advantages	Disadvantages
• Incorporates components, interactions, mechanisms of action • Same toxicological data requirements as for individual data • Site exposure data not needed • Same risk value calculations as for individual chemicals • Utilizes available data on similar mixture	• Original mixture composition varies • Mixture composition changes in the environment • Lack of toxicity data for similar mixtures • Professional judgment to determine sufficient similarity

Comparative Potency Procedure

A group of similar mixtures is assembled for the comparative potency procedure. The mixtures should have mostly the same makeup but differ in ratios of compounds or contain somewhat different components; the range of compounds in the mixture of concern should be reflected by the similar mixtures chosen. It is assumed in this procedure that the class of compounds represented by the similar mixture and the mixture of concern acts with reasonable expectation of biological similarity. Therefore, a simple linear relationship can be drawn from the similar mixtures potencies and applied to the mixture of concern (U.S. EPA, 2002a). This procedure has been used to assess the genotoxic carcinogenic potential of combustion products (Seed *et al.*, 1995).

In order for the comparative potency procedure to work, the mixture of concern must have some *in vitro* data or short term *in vivo* data suitable for the endpoint or mode of action in question (e.g., Ames assay or mouse skin painting assay can be used for a suspected genotoxic mixture). Similar short term data must be available for the similar mixtures and at least one of the similar mixtures must have chronic *in vivo* data available (Seed *et al.*, 1995; U.S. EPA, 2000). The human or animal chronic studies used should concur with the route of human exposure and the expected endpoints (e.g., genotoxic carcinogenicity). For the diesel emission genotoxicity assessment, mouse skin painting assays served as the short term assay; human lung cancer epidemiological data were available for roofing tar and coke oven emissions, which served as the similar mixtures (Schoeny and Margoshes, 1989).

It is assumed that the shape of the dose-response curves will be the same between these chemicals, regardless of the assay or toxicity endpoint (short term bioassay or carcinogenicity). The mixture having chronic data becomes the reference mixture having a toxicity of 1.0 and the similar mixtures and mixture of concern are referred to as having a proportion of the reference mixture's toxicity (U.S. EPA, 2000). It is assumed that if the mixture of concern is twice as potent as the reference mixture in the short-term assay, it will be twice as potent as a human toxicant (Schoeny and Margoshes, 1989). If these potency ratios are maintained across more than one type of short term assay, then more confidence will be placed in this comparative potency assessment (U.S. EPA, 2000). The potency ratio from the most sensitive endpoint should then be used to calculate a human potency estimate for the mixture of concern (Schoeny and Margoshes, 1989).

As with the sufficiently similar mixture procedure, comparative potency allows the use of existing and available toxicological information (Table 3). Comparative potency has also been heralded as a method that allows complex mixtures to be treated as such, not "an impossible sum of toxicants". However, this procedure has only been used to evaluate genotoxic carcinogenicity with assumed linear non-threshold relationships (Schoeny and Margoshes,

1989). This procedure only considers one health effect and cannot be used to evaluate additional effects the mixture may have (Krishnan *et al.*, 1997; NORA, 2004). The outcome of comparative potency is a reference mixture with a potency of exactly 1.0 and the mixture of concern with some ratio of that potency; data variation is lost. These potencies are not real numbers and can only be used as ratios in further calculations. Again, suitable toxicity data are limited; short-term assays must be performed similarly for each of the similar mixtures and the mixture of concern (U.S. EPA, 2000).

Considerable professional judgment must be applied with the comparative potency procedure and support of the assumptions must be well described in the risk assessment document. Similarity of the mixtures must be supported by proving a common mode of action, consistent parallel results over multiple short term assays or component structural similarity (U.S. EPA, 2000). Knowledge of the major mixture contents and their toxicities is vital. Assay applicability to the expected human endpoint is another portion of the assessment requiring professional judgment; validated assays are the best choice (Schoeny and Margoshes, 1989). Further, extrapolations from short term to chronic, *in vitro* to *in vivo*, or animal to human must be substantiated and uncertainty analyzed (U.S. EPA, 2000).

Table 3. Advantages and Disadvantages of the Comparative Potency Procedure

Advantages	Disadvantages
• Incorporates components, interactions, mechanisms of action • Site exposure data not needed • Utilizes available data on similar mixtures	• Original mixture composition varies • Mixture composition changes in the environment • Lack of toxicity data for similar mixtures • Has been applied only to genotoxic carcinogens • Evaluates single health effect • Comparative potency ratio is not a real number; it's a unitless ratio • Professional judgment to determine similarity

Risk Calculation Procedures for Mixture Component Approaches

Mixture component approaches were intended for relatively simple mixtures comprised of ten or so chemicals. Component approaches may be broken into dose addition, response addition and interaction. Dose addition is used for toxicologically similar compounds while response addition should be used for toxicologically different compounds (U.S. EPA, 2000). Ideally, dose addition would be used when all compounds followed the same mode of action in the same target organ; for example, rat studies using four nephrotoxins with the same mode of action result in dose additive effects. Response addition would then be used for all other mixtures whose components follow different modes; rat studies using nine compounds with unrelated target sites or four nephrotoxins with different modes of action displayed response additivity (Cassee *et al.*, 1998). However, there is no consensus or guidance on how strict the requirements for choosing dose or response addition have to be before applying either one (Altenburger *et al.*, 2003; Price *et al.*, 2002). Environmental mixtures of components with the same mode of action are the exception, not the rule (Cassee *et al.*, 1998). If it is unclear which should be used, both dose and response addition may be calculated and the results would represent the range of expected effects if no interaction occurs (Altenburger *et al.*, 2003).

When neither dose nor response addition assumptions appear to fit, further investigation and the use of interaction approaches may be warranted (U.S. EPA, 2000). Interaction approaches factor antagonistic or synergistic responses into the risk description of the mixture (Groten, 2000). All component approaches provide a "snapshot" of the complicated dose-response processes as toxicity will vary relative to proportions of components, the endpoint being assessed, exposure route and duration (U.S. EPA, 2000). The three approaches will be discussed below.

For a true assessment of mixture components using any of the three approaches, a dose-response curve for each chemical is needed (Price *et al.*, 2002). This is due to the definition of "no interaction" and the dependent definitions of synergism and antagonism

discussed previously (U.S. EPA, 2000). Dose-response curves for individual components define what is expected for "no interaction" or simple dose additivity; it is impossible to define antagonism and synergism without knowing what to expect from "no interaction" (Groten, 2000). Commonly, an isobologram is developed from the dose-response curves of two chemicals to predict dose addition. Percentages of response for a single effect are noted for both chemicals; these percentages are graphed together, with one chemical's dose on the x-axis and the other on the y-axis. The isobole is the line formed where the percentage responses of both chemicals intersect (Figure 1) (Altenburger *et al.*, 2003; Dressler *et al.*, 1999). Statistical methods are used to determine if there is a significant deviation from dose addition (Schwartz *et al.*, 1995; U.S. EPA, 2000). Unfortunately, the usefulness of isobole graphs is limited to binary mixtures (Altenburger *et al.*, 2003; Dressler *et al.*, 1999). Computer programs have been developed to plot isoboles and calculate expected results for "no interaction" (dose addition) and "zero interaction" (response addition) between chemicals. CombiTool© (Institute of Molecular Biotechnology, Jena, Germany) can incorporate a fixed dose of a third chemical in its expected results calculations. The dose of the third chemical can be varied systematically to determine its effects on the other two chemicals (Dressler *et al.*, 1999; Feron *et al.*, 2002).

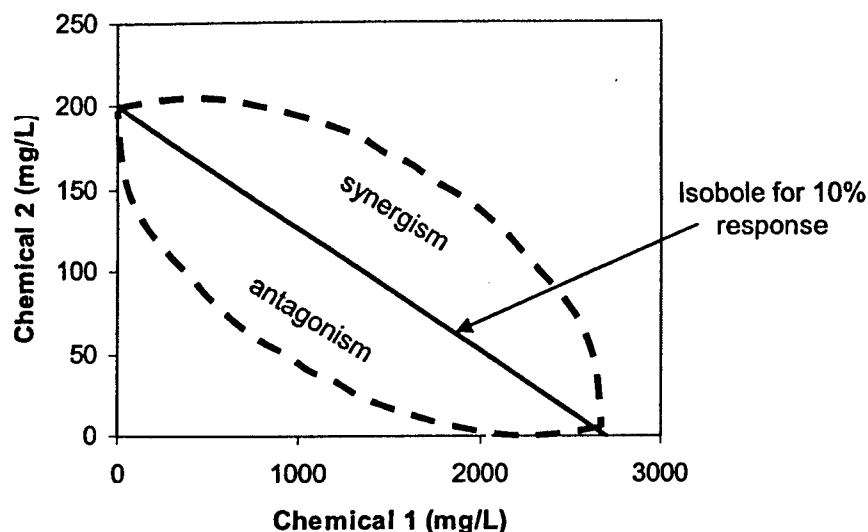


Figure 1. Example isobole for two hypothetical chemicals. The isobole is the line formed by the intersection of 10% responses of both chemicals. Synergistic reactions plot above the isobole and within the dotted line. Antagonistic reactions plot below the isobole and within the dotted line. Adapted from U.S. EPA (2000).

Response Addition Approach

Response addition is also known as Bliss independence; Bliss's description of independent action dates from 1939 (U.S. EPA, 1986). The approach assumes that for two chemicals, when the mode of action is not the same, interaction does not occur (Craig *et al.*, 1999) or interaction will be insignificant at environmental exposure levels (U.S. EPA, 2000). Therefore, toxicity of the mixture can be predicted from individual component responses and the tolerance correlation (ATSDR, 2001a). Tolerance correlations are valued -1, 0 or +1 and provide an estimate of the sensitivity of the exposed population to the two chemicals. A positive correlation would indicate that the receptors most susceptible to the first chemical are also most susceptible to the second. A negative correlation indicates that the population most susceptible to chemical A is least susceptible to chemical B. At zero correlation, no relationship is seen

between the chemicals for receptor susceptibility (U.S. EPA, 2000); the susceptibilities are statistically independent (ATSDR, 2001a).

A good dose-response curve for each component is necessary for calculating response addition; the data, cancer or non-cancer, should be expressed as percent responding (Teuschler *et al.*, 2001). Site exposure data for the components and tolerance correlation estimates are also needed. The probability of response (percent responding) for each component is estimated from the chemical's dose-response curve. If the tolerance correlation is +1, the estimated probability of response for the mixture is equal to the highest of the individual probabilities. The same proportion of the population is expected to be susceptible to both chemicals, so the most toxic chemical will have an effect first; protecting for the most toxic of the chemicals will prevent toxicity from the other chemicals (U.S. EPA, 2000). If the tolerance correlation is -1, different portions of the population are expected to be impacted, so the probabilities of effect for the individual chemicals are added together; this is the most conservative assumption and may be used if tolerance data are missing (Feron and Groten, 2002; Mumtaz *et al.*, 1994). With a zero tolerance correlation, the chemicals act independently by the following relationship (Equation 1):

$$p_{mix} = p_1 + p_2(1 - p_1) \quad \text{(Equation 1)}$$

where p_{mix} , p_1 and p_2 are the probability of effect for the mixture and the first and second chemicals at site concentrations, respectively (U.S. EPA, 2000). This equation reflects the idea that some of the population would respond to the first chemical and some of the remaining population would respond to the second chemical. The zero tolerance correlation equation is used for cancer risk assessment. At very low environmental concentrations and probabilities of effect, Equation 1 effectively becomes simple response addition, the same approach used for -1 tolerance correlations (ATSDR, 2001a).

An advantage of response addition is ease of calculation (U.S. EPA, 2000) (Table 4). It has been used successfully for U.S. EPA Total Cancer Risk of genotoxic mixtures; the probability value for each component is derived from the upper bound estimate of the low-dose linear dose response curve (ATSDR, 2001a; Kodell *et al.*, 1995; Putzrath, 2000). Response addition with a tolerance correlation of +1 is used by the American Council of Government Industrial Hygienists (ACGIH) for determining threshold limit values (TLVs) for mixtures of chemicals having different effects on independent organs (ATSDR, 2001a).

A disadvantage of this approach is the need for population data to determine the tolerance correlation; this requirement limits the use of response addition in risk assessment (U.S. EPA, 2000). Toxicity data on mode of action, to establish independence, and dose-response data to derive the effect probability at site exposure levels are also insufficient for many chemicals (Teuschler and Hertzberg, 1995). Additionally, tolerance correlation works well for two chemicals but its applicability for three or more is limited; three chemicals can not all be negatively correlated with each other (U.S. EPA, 2000). Furthermore, most mixtures found in the environment are made up of multiple chemicals; use of component approaches assumes the toxicologically relevant contaminants are identified and addressed, whereas one or more significant toxins may have gone undetected (Feron and Groten, 2002; Groten, 2000).

Response addition assumptions are highly dependent on exposure levels and potentially dependent on the route of exposure. At low doses typical in the environment, two chemicals may have independent effects on the same organ due to different modes of action. At high doses, increased toxicity on the two response systems may lead to physiological interactions and additional toxicity not suspected from responses at lower doses. For this reason, response addition should only be used at or near individual component no observed adverse effect levels (NOAELs). If data are not sufficient to conclude that toxicity will remain independent at environmental levels, dose addition would be the conservative approach to use (U.S. EPA, 2000). If the independence assumption is in error, response additivity will underestimate the

risk from the mixture (Mumtaz *et al.*, 1994; U.S. EPA, 2000). During the risk characterization phase, uncertainty regarding poorly studied components that could dominate the response addition probability value should be thoroughly discussed (U.S. EPA, 2000).

Table 4. Advantages and Disadvantages of the Response Addition Approach

Advantages	Disadvantages
• Easy to calculate • Frequently used for genotoxic cancer risk assuming zero tolerance interaction	• Data limited to determine tolerance correlation • Data insufficient to determine independence and dose-response curve for some chemicals • Best suited to binary mixtures; Assumption of tolerance correlation allows use for larger mixtures • Assumes all toxic components addressed • Applicable near NOAELs, extrapolation to higher doses inappropriate • Underestimates risk if independence assumption erroneous

Dose Addition Approaches

Dose addition approaches assume that since all chemicals present in the mixture have the same mode of action, they can be considered a dilute or concentrated solution of one chief component (U.S. EPA, 2000) and their concentrations can be added together to predict the toxic effect (Mumtaz *et al.*, 1997; U.S. EPA, 1986). Also known as Loewe additivity, the components of dose addition are assumed to have parallel dose-response curves, similar kinetics and the same toxicodynamic action (ATSDR, 2001a; U.S. EPA, 2000). Additivity procedures are expected to provide risk estimates that are neutral, meaning neither conservative nor underprotective, if used for components that produce similar effects in the same organ (U.S. EPA, 1986).

An advantage of dose addition or additivity is that the approach predicts well even when the assumptions above are not strictly valid. Additivity overpredicts the toxicity of dissimilar chemicals (i.e., different modes of action). However, studies using mixtures of similar chemicals (i.e., same mode of action) along with dissimilar components have shown that additivity predicts the toxicity reasonably well. Another advantage of dose addition is the relative availability of the toxicological data needed to perform the assessment (U.S. EPA, 2000); individual component data are frequently available from EPA's IRIS (U.S. EPA, 2004a) or ATSDR toxicological profiles (ATSDR, 2001a).

Disadvantages of dose addition include the fact that potential interactions between components are not considered (ATSDR, 2001a; NORA, 2004). For this reason, qualitative interaction evidence should be sought throughout the dose addition assessment to offset this deficit (Mumtaz *et al.*, 1997). From a risk standpoint, however, additivity and antagonism are found most frequently in binary mixtures; when synergism occurs, it is generally within a factor of 2 of the additivity-predicted response, indicating limited cause for concern (Dyer, 2002). Further, dose addition calculations incorporate environmental exposure levels of chemicals, which are generally near the NOAELs, where interactive effects are not expected to occur often (U.S. EPA, 2000). However, sufficient data necessary to check the assumption of parallel dose-response curves are unavailable for some chemicals and information on pharmacokinetics and mode of action are generally even scarcer (Teuschler and Hertzberg, 1995).

Hazard Index Procedure

The hazard index (HI) is the most well-known dose additive procedure to account for non-cancer mixture risks. Summing of hazard quotients (HQs) from chemicals with the same target organ to get a mixture HI is recommended in RAGS A (U.S. EPA, 1989) and so is the default procedure in Superfund risk assessments. The procedure is also used by the ACGIH and the Occupational Safety and Health Administration (OSHA) (ATSDR, 2001a). The

Department of Energy (DOE) calculates emergency airborne exposures by the HI procedure (Craig *et al.*, 1999).

The HI approach assumes that the HQs of individual chemicals can be summed to calculate expected risk for a site specific mixture. HQs are the quotient of the environmental exposure concentration divided by a "safe" exposure level. The "safe" level in the denominator is determined by the regulatory reason for calculating site risk; at Superfund and Resource Conservation and Recovery Act (RCRA) sites, the denominator is most often an RfD or RfC or sometimes an acceptable daily intake (U.S. EPA, 1986; 2000). ACGIH and OSHA divide the eight-hour time weighted average exposure concentration for individual chemicals by the chemical's threshold limit value (TLV) or permissible exposure level (PEL) to derive their HQs, respectively (ATSDR, 2001a). For emergency planning purposes, the DoE recommends dividing the maximum expected airborne concentration of each component by its temporary emergency exposure limit (TEEL) (Craig *et al.*, 1999). The HI approach attempts to approximate the toxicity of the mixture if the mixture were to be tested in entirety (U.S. EPA, 2000).

RAGS A recommends calculating a screening HI, which includes all chemicals regardless of target organ (U.S. EPA, 1989). This screening HI is common practice (ATSDR, 2001a; Mumtaz *et al.*, 1997; U.S. EPA, 2004b). If the result of the screening calculation is less than 1 for the entire mixture and there is no evidence of interaction between chemicals in the literature, then the mixtures assessment ends. If the screening result is near or greater than one, target organ HIs are then calculated (Mumtaz *et al.*, 1997).

By the definition of dose addition, a HI should be calculated for each mode of action within a target organ when the screening HI exceeds 1. In reality, mixture risk assessments include one HI for each target organ or, when data are available, each endpoint for every route of exposure (U.S. EPA, 2000). A systematic process can be used to determine which organs or endpoints the chemical affects. The U.S. Army Center for Health Promotion and Preventative

Medicine (USACHPPM) is developing a hierarchical chart of 14 target systems (e.g., skeletal, nervous) divided into 76 target organs (e.g., bone, ligaments and cartilage, connective tissue) to help guide placement of chemical found at a site into the appropriate HI calculations (Johnson, personal communication, 2004). Similarly, the DoE uses health code numbers such as those published in Patty's Industrial Hygiene and Toxicology (John Wiley and Sons, NY) to sort chemicals by target organ and mode of action (Craig *et al.*, 1999).

If the result from a target organ HI is greater than one and mechanistic data are available, then mode of action HIs should be calculated (Mumtaz *et al.*, 1997). If a final target organ HI or a mode of action HI exceeds 1, more investigation of site specific toxicity or remediation of the site is indicated (U.S. EPA, 2000). HIs can sometimes be used to rank sites within a property or installation, indicating which is of greater toxicological concern; site ranking is more defensible if the sites have similar types of contamination. The disadvantage of this approach is that a HI of 4 does not make the site twice as bad as a site with an HI of 2 (Mumtaz *et al.*, 1994). The HI is a unitless sum of ratios with order of magnitude precision and does not reflect probability of toxic effect (U.S. EPA, 2000). The site with a HI of 4 would instead be considered as more urgent in its need for further study or remediation (Mumtaz *et al.*, 1994). Further study may be warranted as a single HQ may drive the HI over 1 (ILSI, 1999). An excessive HI may be due to some components of the mixture having RfDs based on large uncertainty factors. HI values used in risk decisions should be accompanied by complete characterization of the uncertainty, including poor quality of information and data gaps (U.S. EPA, 2000). Uncertainties for a mixture of just a few components can be substantial (U.S. EPA, 2002c).

An advantage of the HI procedure is its familiarity and ease of use (U.S. EPA, 2000) (Table 5). The data required for a mixtures HI are the same as required for individual chemical risk calculations (Putzrath, 2000). Unfortunately, the procedure is just as easily misused. Target organ HIs are frequently calculated without regard to available mode of action data;

these HIs are then utilized in decision making for the site, even though mixture risk from a common mode of action may not exist (ATSDR, 2001a). The HI procedure is also easy to use due to the ready availability of appropriate toxicity data such as RfDs and RfCs for many chemicals (ILSI, 1999); unfortunately, mode of action and dose-response data to verify the applicability of dose addition are not always available (Teuschler and Hertzberg, 1995; Viau, 2002). Therefore, the requirement of applying dose addition to chemicals that act by the same mode of action is relaxed; HIs are calculated for mixture components that affect the same target organ or system (U.S. EPA, 2000).

Use of RfDs and RfCs in HI calculations has both advantage and disadvantages. EPA reference values are products of extensive scientific peer review prior to posting on IRIS, where they are readily available (U.S. EPA, 2004a). These values incorporate uncertainty factors to offset inadequacies in our knowledge of the chemical's toxicity (ILSI, 1999). However, these uncertainty factors adjust the reference values down orders of magnitude, resulting in HQs that are sometimes seen as overly conservative (Carpy *et al.*, 2000; Mumtaz *et al.*, 1997). This problem is compounded by the fact that the point of departure for an RfD is the NOAEL or lowest observed adverse effect level (LOAEL) of the most sensitive effect. Secondary effects or alternate sites of toxicity are ignored in the final RfD value. Frequently, target organ HIs are calculated for all applicable systems where mixtures toxicity could occur but are determined using reference values based on a different organ system (Mumtaz *et al.*, 1997; U.S. EPA, 2000).

Reference value target organ specificity can be offset by the calculation of a comparable target organ toxicity dose or concentration (TTD or TTC) used in the place of a RfD or RfC except when calculating mixtures toxicity for the target organ on which the RfD was based (U.S. EPA, 2000). TTDs are calculated using the EPA RfD derivation process, except the NOAEL or LOAEL for the target organ is used, instead of the most sensitive endpoint overall. Applicable uncertainty factors are assigned to the target organ NOAEL. An uncertainty factor for database

deficiency is only used if the deficit is applicable to the target organ; lack of reproductive and developmental studies should not add an uncertainty factor to a liver TTD (Mumtaz *et al.*, 1997). TTDs attempt to account for the fact that most chemicals have more than one site of toxic action and that these secondary effects occur only at higher exposure levels; this approach is being promoted in the "Interaction Profiles" under development by the ATSDR (e.g., ATSDR, 2001b, 2002). However, the TTD process requires additional chemical toxicity data for all important target organs and a peer review to validate the TTD levels. For chemicals already assessed by the ATSDR, target organ data can be found in their "Toxicological Profiles"; if sufficient target organ data can not be found, the RfD is used as a conservative substitute for the chemical in the mixture target organ HI (ATSDR, 2001a).

Table 5. Advantages and Disadvantages of the Hazard Index Procedure

Advantages	Disadvantages
• Predicts well for similar and dissimilar mixtures • Toxicity data readily available • Familiar and easy to use • RfDs and RfCs are peer reviewed and incorporate uncertainty factors • TTDs relevant to target organ HIs	• Interactions not considered • Data insufficient to verify dose-response curves and mode of action • Reference values are target organ specific • TTDs require additional data and peer review

Fraction Surrogate Procedures

Fractionated surrogate procedures are specialized applications of the hazard index procedure. Fractionated assessments assume that the toxicity of a chemically defined portion (fraction) of the whole mixture can be determined by the toxicity of a surrogate. When the surrogate is the most toxic compound within the fraction, the entire fraction is assumed to have the greatest potential toxicity (Hutcheson *et al.*, 1996). These procedures are in use for total

petroleum hydrocarbon (TPH) contamination and have been proposed for PAHs to account for the nongenotoxic congeners (Brown *et al.*, 1999).

Three main TPH surrogate procedures have been developed for use nationally. In all three, cancer risk is determined first by evaluating site exposure levels of common carcinogens present in TPH including benzene and specific PAHs (U.S. EPA, 2002d); use of the indicator compound approach for carcinogenic risk is expected to account for the most toxic compounds and be protective of human health (Hutcheson *et al.*, 1996). Each procedure has utilized the RfD derivation process to determine noncancer risk values for each of the fractions based on toxicity information for surrogate components within the fractions (U.S. EPA, 2002d).

The Massachusetts Department of Environmental Protection (MA DEP) was the first to introduce a fraction surrogate procedure for TPH in 1994 (MA DEP, 2003); their procedure has been adapted by Alaska (Reed and Sterner, 2004). MA DEP developed two gas chromatography/flame ionization detector analytical methods, volatile petroleum hydrocarbons and extractable petroleum hydrocarbons, to split TPH soil contamination into six total aromatic and aliphatic fractions based on absolute carbon number (e.g., aliphatic C₅-C₈ fraction). These fraction exposure concentrations measured from site soil are divided by the MA DEP toxicity values based on a surrogate compound to develop an HQ for each fraction. MA DEP surrogate compounds were chosen to be protective for all components in the fraction; most surrogates considered were already well characterized and many had RfDs and RfCs. HQs are summed for each exposure route HI; measured site concentrations are extrapolated to predicted exposure concentrations for some routes using simple fate and transport models (MA DEP, 2003).

Concurrently with the development of the MA DEP procedure, a national ad hoc group representing the DoD, petroleum industry and consultants, developed their own analytical method and risk values for the resulting fractions (MA DEP, 2003). The Total Petroleum Hydrocarbon Criteria Working Group (TPHCWG) developed a gas chromatography method

known as the direct method to divide TPH soil contamination into 13 total aliphatic and aromatic fractions based on effective carbon atom number. The effective carbon number index is related to the boiling point of the chemical and helps separate chemicals based on environmental transport properties. The site exposure concentrations for each fraction are divided by the TPHCWG risk criterion for the fraction. Instead of focusing on chemicals already having EPA RfDs and RfCs, the TPHCWG reviewed available toxicity studies for as many fraction components as possible before developing reasonably conservative risk values representative of the fraction. Like the MA DEP procedure, a HI is calculated for each route of exposure by summing the HQs for each fraction; again exposure concentrations for some routes are predicted using simple environmental fate models (Vorhees *et al.*, 1999). Several states have accepted or adapted the TPHCWG procedure including Louisiana, Texas, Oregon and Washington (Reed and Sterner, 2004). This procedure has also been used successfully in the United Kingdom (Clay and Harris, 2002) while Canada has its own similar procedure (Sevigny *et al.*, 2003).

Recently the U.S. EPA has proposed its own Provisional Peer Reviewed Toxicity Values (PPRTV) for TPH contamination at Superfund sites; the reference values were made available to the public in 2004. The PPRTVs combine many of the characteristics of both the MA DEP and TPHCWG fractions and risk criteria but more closely resemble MA DEP's procedure. The PPRTVs require MA DEP's or very similar analytical methods and assess the non-cancer health risks of benzene, toluene, ethylbenzene and xylenes separately from the fractions, similar to MA DEP (U.S. EPA, 2002d). The PPRTVs will be used in the future at Superfund sites with TPH contamination (U.S. EPA, 2003).

The primary advantage of fraction surrogate procedures is their usefulness for highly complex and variable mixtures (Table 6). TPH contamination is widespread; contamination consisting of only TPH accounted for 50% of Massachusetts' hazardous waste sites prior to the development of their fraction approach (Hutcheson *et al.*, 1996). Yet TPH is not one entity

suitable for whole mixtures toxicity testing. This mixture consists of hundreds of aliphatic and aromatic hydrocarbons that differ depending on the product (e.g., gasoline, diesel) which was released into the environment, the crude from which the product was refined and the transformation processes in the environment during weathering (Hutcheson *et al.*, 1996; U.S. EPA, 2002d; Vorhees and Butler, 1999). Component based dose or response addition is not suitable for a TPH site due to the analytical and calculation workload necessary to attempt adding toxicity for hundreds of chemicals, many of which do not have sufficient health effects data (Hutcheson *et al.*, 1996).

Another advantage is that fraction surrogate assessments assign toxicity based on the contaminants present at the site. Health effects associated with TPH contamination do not correlate with analytical measures of total TPH because components differ from site to site (Hutcheson *et al.*, 1996). Aside from total TPH regulations, other previous methods of dealing with TPH cleanup levels were to use site indicator compound levels, such as benzene (MA DEP, 2002). Volatile toxins like benzene are often absent from older sites, yet considerable amounts of more persistent compounds remain. Fraction surrogate procedures provide practical risk-based assessments for non-carcinogenic endpoints (U.S. EPA, 2002d).

The MA DEP and TPHCWG procedures and their adaptations have been demonstrated and utilized for setting cleanup levels at multiple sites across the country. In a side-by-side comparison of the two procedures, the risk based screening levels were found to be comparable (Reed *et al.*, 2003). Since the EPA's PPRTVs are similar to the MA DEP risk values, it is expected that the EPA procedure will also produce comparable cleanup levels. These procedures could be adapted for TPH ecological risk assessment by using one of the fractionation analytical methods and toxicological risk criteria based on sensitive ecological endpoints (Twerdok, 1999).

A major disadvantage of the fractionation procedures is the analytical cost versus the costs involved with measuring a few indicator compounds and total TPH; in addition, there are

relatively few analytical laboratories which perform the fractionation procedures (Hutcheson *et al.*, 1996). Chemical extraction procedures prior to analysis do not reflect bioavailability of the TPH components in the environmental setting (Reed and Sterner, 2004). Further, some uncertainty and variability are inherent in the process of dividing the fractions prior to quantitation (Hutcheson *et al.*, 1996).

Like all dose addition approaches, fraction surrogate procedures do not consider interaction, which could result in over- or underestimated risk (Hutcheson *et al.*, 1996). However, the dose addition assumption potentially overestimates risk. Fractions are composed of chemically similar compounds; the toxicological similarity is unknown so dose addition is a conservative assumption. The different fractions, however, are dissimilar chemically and are even less likely to have toxicological similarity. The applicability of response independence would be difficult if not impossible to confirm for such a complex situation, so again dose addition is the health protective assumption (U.S. EPA, 2000). MA DEP has developed conservative soil screening levels for individual fractions to be used with initial site assessments (MA DEP, 2003). Fraction risk is not summed; therefore response addition with a tolerance correlation of +1 is assumed.

Table 6. Advantages and Disadvantages of the Fraction Surrogate Procedures

Advantages	Disadvantages
• Fraction risk criteria readily available • Provides health-based assessment means for highly complex and variable mixtures • Accepted for use in several locations and now at Superfund sites • Comparable results between procedures • Potentially adaptable to ecological risk	• Interactions not considered • Data insufficient to verify dose addition assumptions • Relatively high analytical cost • Chemical extraction ignores bioavailability • Uncertainty and variability in separating fractions

Relative Potency Factor/Toxic Equivalency Factor Procedure

The relative potency factor (RPF)/toxic equivalency factor (TEF) procedure is a dose addition approach suitable for specific contaminants. In this procedure, the toxicity of an index chemical with an adequate toxicity database is used to infer the health effects of similar chemicals with the same mode of action (U.S. EPA, 2000). This procedure depends on the additive assumptions: individual components act through the same mode of action, their dose-response curves are parallel and the resulting effects are of combination exposures that are completely additive (NORA, 2004).

RPF and TEF are essentially the same entities. The U.S. EPA considers TEFs and RPFs to convey different certainties behind the same basic approach. TEFs are special RPFs in that they are more scientifically defensible. TEFs apply to all health endpoints resulting from all exposure routes and durations. More data are available for TEF development and the mode of action is more certain. Few chemical classes are expected to meet the data criteria for TEF development. RPFs are limited in their applicability and may be based on fewer data. Less may be known about the mode of action or the components may be suspected of having toxicities not mediated by the receptor on which the RPF is based (U.S. EPA, 2000). RPF use may be restricted to one route of exposure and extrapolation to other routes prohibited (U.S. EPA, 2004b). TEFs have been developed for dioxins, furans and 13 dioxin-like PCBs; RPFs have been developed for carcinogenic PAHs. Because the confidence in mechanistic knowledge behind TEFs and RPFs differs, scientific and regulatory acceptance will also differ (U.S. EPA, 2000). The distinction between RPF and TEF often is not clear outside EPA publications; TEF is the term commonly used in literature and at sites (Pohl *et al.*, 1997; Reeves *et al.*, 2001) and will be the term used in this review unless the point is specific to RPF.

TEFs are developed for persistent chemically related compounds occurring in environmental mixtures that are thought to produce toxicity through a common receptor. Within this well-defined group of congeners, there must be an index compound with a toxicity database

adequate for creating a quantitative dose-response assessment for a specific exposure route. The index compound for dioxins and dioxin-like compounds including the 13 co-planar PCBs is 2,3,7,8-tetrachlorodibenzo-[p]-dioxin (2,3,7,8-TCDD or TCDD), based on carcinogenicity mediated through aryl hydrocarbon (Ah) receptor activity. Similarly, the index compound for carcinogenic PAHs is benzo[a]pyrene due to its binding with the Ah receptor (U.S. EPA, 2000). Minimal data requirements for all other congeners are *in vitro* or short term *in vivo* assays in common (i.e., same species, duration and endpoint) with the index compound (ILSI, 1999; U.S. EPA, 2000). Of course, longer term *in vivo* study results are preferred as kinetics are not a factor *in vitro* and, *in vivo*, steady state is often not achieved in a few days when testing the persistent chemicals for which the TEF procedure was designed (Carpenter *et al.*, 2002; Putzrath, 1997; Seed *et al.*, 1995). Toxicological outcomes should be shared by the congeners and attributable to the mode of action. Inconsistencies with the mode of action should be well documented (U.S. EPA, 2000).

A TEF is the numerical ratio of the toxicity value of the congener over the toxicity value of the index compound (U.S. EPA, 2000). The toxicity values used may be NOAELs, LOAELs or benchmark doses (BMDs) from bioassays that the congeners and index chemical have in common (U.S. EPA, 2000; 2002c). The choice of toxicity value has been shown to affect the TEF by more than a factor of two; the best available toxicity value, generally a NOAEL or BMD, must be used (ILSI, 1999; Seed *et al.*, 1995). Multiple endpoints are often compared to calculate a range of TEFs (Kleinjans, 2003; Reeves *et al.*, 2001). Conservatively, the upper bound confidence value for this range is generally assigned as the TEF (Seed *et al.*, 1995). TEF ratios must be established through scientific consensus and peer review prior to use in an assessment (U.S. EPA, 2000).

TEF ratios are used by multiplying the amount of the congener present at the site by the TEF to determine an adjusted amount of the index compound, known as a toxic equivalent (TE or TEQ). Congener TEQs are summed to form the equivalent concentration of the index

compound. The index compound's dose-response curve is then used to estimate the effect of the summed TEQs (U.S. EPA, 2000; 2002c). Unlike the unitless HI ratio, this equivalency is an estimated concentration of a chemical which is associated with specific toxicity (e.g., pg/kg/day TCDD) and can be loosely considered as that chemical throughout risk decision making (Safe, 1994; Schoeny and Margoshes, 1989; Teuschler and Hertzberg, 1995).

The main advantage of the TEF procedure is its legitimacy when used for common receptor mediated effects (Table 7). The procedure provides a practical solution for dealing with understudied congeners. It prevents underestimation of health hazard by allowing risks to be calculated for congeners without sufficient data for an RfD; without the TEF procedure, only congeners with RfDs would figure into risk calculations and the understudied congeners would be noted as part of the uncertainty write-up (van den Berg *et al.*, 2000). Pesticides are often applied in combination and succession. Using TEFs, the U.S. EPA normalizes concentrations of pesticides working by the same mechanism to predict potential total exposure levels in the environment (U.S. EPA, 2002c).

Disadvantages of the TEF procedure are similar to those of any dose additive approach. Comprehensive knowledge of the mixture composition and amounts of individual components are required to calculate mixtures risk for the site. Interaction between components is not significant and additivity is assumed to be absolute. This assumption may not hold even at low concentrations; competitive binding between congeners can produce interactions and metabolic pathways can become saturated (Borgert *et al.*, 2003; Kleinjans, 2003; Krishnan *et al.*, 1997; NORA, 2004; Pohl *et al.*, 1997). Mixtures of most dioxin-like compounds have been shown to be additive, but others result in antagonism and synergism, depending on receptor species and endpoint tested. Non-dioxin-like PCBs can antagonize TCDD effects (ATSDR, 2001a; Carpenter *et al.*, 2002; Rao and Unger, 1995; Safe, 1994). Uncertainties for all these assumptions should be thoroughly characterized when setting a TEF ratio (U.S. EPA, 2000).

TEF approaches give a false sense of toxicological knowledge. TEFs do not have precise numerical values but are order of magnitude estimates (Putzrath, 1997). Only one mode of action and one toxicological endpoint is examined for all congeners; the resulting TEF is expected to be protective for all endpoints (U.S. EPA, 2000). The RPF for PAHs and TEFs for dioxin-like congeners are based on carcinogenic effects related to Ah receptor binding; other sensitive effects such as adverse reproductive endpoints have also been associated with the Ah receptor. Many dioxin-like and non-dioxin-like PCBs have effects apparently not mediated by the Ah receptor pathway, including neurological deficits and endocrine disruption (Carpenter *et al.*, 2002). The ATSDR considers the TEF approach unsuitable for PCBs, due to their non-Ah receptor mediated toxicity and the fact that only 13 of 209 congeners have dioxin-like effects. Both the ATSDR and USACHPPM consider TEFs suitable for PAHs carcinogenic risk assessment due to substantiating human epidemiological evidence, but not applicable for non-cancer effects (Johnson, personal communication, 2004; Pohl *et al.*, 1997). TEFs may discourage investigation of secondary effects from environmental mixtures (Putzrath, 1997).

Further, TEFs are considered interim assessments until a more definite assessment is possible, such as RfDs or other suitable values for each component being set to calculate HIs for the mixture (U.S. EPA, 2000). However, this interim status means that revisions to the TEFs continue on a regular basis, allowing them to remain up-to-date (U.S. EPA, 2004b). It is possible to combine the HI procedure for well-studied components with the TEF procedure for relatively unstudied congeners, or, for PCBs, to combine the TEF approach for the 13 dioxin-like PCB congeners with the total PCB slope factor for the remaining PCB concentration (Putzrath, 1997; U.S. EPA, 2000). This last option is particularly viable as these 13 dioxin-like PCBs are highly persistent in environmental media and food chains (Safe, 1994).

TEFs focus on single receptor endpoints. However, the TEF procedure will not work for all receptor mediated effects. For example, the estrogen receptor is not one receptor; it has multiple forms activated in different ways. Xenoestrogens are structurally different, not true

congeners. The kinetics of estrogen mimics vary widely, making *in vitro* experiments difficult to extrapolate (Borgert *et al.*, 2003). The restriction to single receptor effects from structurally related chemicals must be adhered to in order to use the TEF procedure (U.S. EPA, 2000).

Table 7. Advantages and Disadvantages of the TEF Procedure

Advantages	Disadvantages
• Predicts well for chemicals with same mode of action • Uses available toxicity data, <i>in vivo</i> and <i>in vitro</i> • Toxic equivalents stated in concentration units of index compound • Predicts potential cumulative pesticide exposure • Allows risk quantitation for congeners without sufficient toxicity databases	• Interactions not incorporated • Data insufficient to verify dose-response curves and mode of action • Needs detailed knowledge of mixture components and concentrations • Ignores additional sensitive endpoints and modes of action • Interim assessment method • Does not work for all receptor types

Interaction Approaches

The main disadvantage of response addition and dose addition is that neither approach accounts for chemical to chemical interactions outside of documenting them as uncertainty. Interactions approaches attempt to incorporate available antagonism and synergism data into mixtures risk assessment in a systematic way. However, these approaches are not used in the majority of mixtures risk assessments; regulatory agencies are not requiring them as the approaches are data and resource intensive (U.S. EPA, 2000).

A major hindrance to interaction assessment is the scarcity and frequent poor quality of interactions data (Mumtaz *et al.*, 1994). Only a relatively limited number of binary interaction studies, toxicity studies of two individual chemicals alone and in combination, have been published (Seed *et al.*, 1995). Most of those studies focus on liver toxicity; rarely are data on additional target organs collected (Simmons, 1995). Extrapolating doses using binary data can be difficult (Feron and Groten, 2002). Often, binary studies involve single, high dosages instead

of multiple doses in lower, environmentally relevant ranges. Frequently the magnitude of interaction, the ratio of observed toxicity to the toxicity expected from additivity, is not documented in literature. The direction of interaction (i.e., synergism or antagonism) and magnitude can change depending on dose or relative proportions of components, sequence of administration, route of exposure and receptor species (Hertzberg and MacDonell, 2002; Seed *et al.*, 1995). Published binary studies sometimes fail to perform statistics, omit statistical procedures from the paper, do not define "no interaction" or end up comparing binary toxicity to individual component toxicity. Without predicting what dose additive toxicity should have been, it is impossible to determine if interactions took place (Borgert *et al.*, 2001; Simmons, 1995).

Interaction studies should be carefully planned. Multiple assays should be involved. *In vitro* studies can detect pharmacodynamic interactions but often *in vivo* studies are necessary to identify kinetic interactions (Borgert *et al.*, 2001; Teuschler *et al.*, 2002). The route of exposure should be environmentally relevant. "No interaction" should be defined as either dose addition or response addition and results compared using appropriate statistical tests to determine significant biological variation from the expected results (Borgert *et al.*, 2001; Hertzberg and MacDonell, 2002). Doses need to be selected carefully. They should reflect environmental exposure concentrations, include the individual component NOAELs and LOAELs, and also be spaced to develop a dose-response curve (Borgert *et al.*, 2001; Cassee *et al.*, 1998; Feron and Groten, 2002; Hertzberg and MacDonell, 2002; Teuschler *et al.*, 2002).

Mixture Weight of Evidence Procedure

The mixture weight of evidence (WOE) procedure is a qualitative or semi-quantitative systematic method of examining possible interactions present in a mixture. This procedure was designed to be used in conjunction with a mixture HI (ATSDR, 2001a; U.S. EPA, 2000). Developed in 1992, it is being used as the primary method of assessment in the ATSDR's Interaction Profiles, currently in draft format (e.g., ATSDR, 2001b; 2002). These profiles

evaluate common environmental co-contaminants for potential interactions using binary data from the literature, the mixture WOE procedure and peer review to validate conclusions (Pohl *et al.*, 2003). It is assumed that binary effects adequately describe the bulk of interactions in a mixture; tertiary and more complex interactions should have minimal impact (ATSDR, 2001a; Seed *et al.*, 1995).

The first step in a mixture WOE evaluation is to review literature information on all binary interactions between mixture components of concern at the site. Data needed include toxicity studies, pharmacokinetic descriptions and mechanistic papers for mixture components and structurally related chemicals. The goal is to determine the "most likely effect" of each chemical on another, forward and reverse (e.g., PCBs on TCE and TCE on PCBs), for each potential target organ system (ATSDR, 2001a; Pohl *et al.*, 1999; U.S. EPA, 2000). For each organ system, a combination is judged qualitatively based on the criteria in Table 8. First a direction is determined, where additive, synergistic and antagonistic effects are "=", ">" and "<", respectively. If a direction can not be determined, a "?" is used and the assessment of the combination ceases. The quality of the interaction or no interaction data is categorized using an alphanumeric system (Table 8). For the two main classifications, mechanistic and significance data are considered best if from the chemical of concern, but may be used from structurally related compounds. Modifying classifications are included to qualify the data used to determine the two main classifications (ATSDR, 2001a; Mumtaz and Durkin, 1992). An example of a mixture WOE classification for PCBs on TCE is ">IIB2" for both hepatic and neurological effects (ATSDR, 2001b).

The mixture WOE classification can be converted to a semi-quantitative evaluation using the direction and weighting factors (Table 8). Each portion of the alphanumeric mixture WOE is multiplied together; the resulting BINWOE (binary weight of evidence) score can range from -1 to -0.05 for antagonist effects, +0.05 to +1 for synergistic effects or equal 0 when effects are

additive. If the mixture WOE was a "?", the BINWOE becomes 0 (Mumtaz and Durkin, 1992; Mumtaz *et al.*, 1994).

Table 8. Mixture Weight of Evidence Alphanumeric Classification and Binary Weight of Evidence Scores*

Direction Classification	Direction of Interaction	Direction Weighting
=	Additive	0
>	Greater than additive	+1
<	Less than additive	-1
?	Indeterminate	0
Alphanumeric Classification	Data Quality Criteria	Quality Weighting
<i>Mechanistic Understanding Criteria</i>		
I	Direct and Unambiguous Mechanistic Data: Mechanism(s) of interaction well characterized; Unambiguous interpretation of interaction	1.0
II	Mechanistic Data on Related Compounds: Mechanism(s) not well characterized for chemicals of concern; Structure-activity relationships (quantitative or qualitative) infer likely mechanisms and direction of interaction	0.71
III	Inadequate or Ambiguous Mechanistic Data: Mechanism(s) of interaction not well characterized; Information on mechanism(s) does not indicate direction of interaction	0.32
<i>Toxicological Significance Criteria</i>		
A	Toxicological significance of interaction directly demonstrated	1.0
B	Toxicological significance of interaction inferred or demonstrated for related chemicals	0.71
C	Toxicological significance of interaction unclear	0.32
<i>Modifying Criteria</i>		
1	Anticipated duration or sequence of exposure	1.0
2	Different duration or sequence of exposure	0.79
a	<i>In vivo</i> data	1.0
b	<i>In vitro</i> data	0.79
i	Anticipated route of exposure	1.0
ii	Different route of exposure	0.79

*Adapted from Mumtaz and Durkin, 1992

In a site risk assessment, the mixture WOE classifications for all interactions are laid out in a matrix box in order to assess the general trend of potential interactions. This trend is used with a HI to help describe the confidence in the HI based on known chemical to chemical interactions. Conclusions on expected interactions may depend on the concentrations of the mixture components found at the site (Mumtaz and Durkin, 1992). At trace levels, competitive interactions are not expected to occur but at higher exposure concentrations there is greater potential for antagonism or synergism (Pohl *et al.*, 1999).

An overall trend may be hard to determine as additive, synergistic and antagonistic effects may be expected to occur at one site, along with several unknown interactions ("?) (ATSDR, 2001a); for this reason, the semi-quantitative procedure was developed. BINWOE scores for individual interactions for a target organ system are added together. This total mixture WOE (WOE_{MIX}) is normalized by dividing by the maximal WOE (WOE_{MAX}), which is the sum of the geometric means of HIs for all pairs of chemicals in the interaction:

$$WOE_{MIX} = \sum BINWOE \quad (\text{Equation 2})$$

$$WOE_{MAX} = \sum \sum_{x \neq y} (HI_x \times HI_y)^{0.5} \quad (\text{Equation 3})$$

where x and y represent different individual chemicals (Mumtaz and Durkin, 1992). The geometric mean is an appropriate measure of central tendency as it minimizes the influence of single values, is applicable for unknown or non-normal distributions and is conservative as the geometric mean is lower than the arithmetic mean (MacDonald *et al.*, 2000; Mumtaz *et al.*, 1994).

$$WOE_N = \frac{WOE_{MIX}}{WOE_{MAX}} \quad (\text{Equation 4})$$

The normalized WOE (WOE_N) represents the level of confidence that an interaction exists at the site exposure levels (Mumtaz and Durkin, 1992). The WOE_N is a smaller number when exposure levels decrease (i.e., HI ratios decrease); interaction is less likely to occur at lower

exposure levels (U.S. EPA, 2000). The normalized WOE modifies the additivity HI (HI_{ADD}) by the following equation (Equation 5):

$$HI_I = HI_{ADD} * UF_I^{WOE_N} \quad (\text{Equation 5})$$

where HI_I is the interaction hazard index for the mixture and UF_I is an interaction uncertainty factor (UF), generally assigned a value of 10 to allow for scaling by the normalized WOE (U.S. EPA, 2000). The WOE_N adjusts the UF up or down depending on predicted synergism or antagonism of the mixture; the resulting UF can range from 0.1 to 10 (ATSDR, 2001a).

One advantage of the mixture WOE method is its adaptability to carcinogenic effects (Table 9). Although developed for non-carcinogenic endpoints, the alphanumeric evaluation and the semi-quantitative WOE_{MIX} (sum of all BINWOEs) are both applied to cancer risk in a qualitative way. If the mixture WOE classification or if the WOE_{MIX} indicates interaction is greater than zero, then a potential health hazard is assumed. This is particularly true when the sum of cancer risks for the mixture (response addition) is near or higher than 1×10^{-4} (ATSDR, 2001a).

The BINWOE method has been shown to be consistent with whole mixture toxicity data. Better agreement was achieved with toxicological results from mixtures of components having the same mechanism than with the results from chemicals with the same target organ but different mechanisms; this result is not surprising as the mixture weight of evidence procedure is based on dose additivity assumptions (Cassee *et al.*, 1998; Mumtaz *et al.*, 1998). Mixture WOE classification and BINWOE values were found to be applied consistently between different groups of risk assessors and toxicologists (Mumtaz *et al.*, 1996). BINWOE scores were found to be similar, although sometimes mixture WOE classifications differed (U.S. EPA, 2000).

However, the BINWOE procedure oversimplifies a complex process. BINWOE values are summed across a target organ system to represent the entire mixture; this value does not change with different proportions of the mixture (ATSDR, 2001a). BINWOE values also do not

take into account the magnitude of interactions between chemicals (Hertzberg and MacDonell, 2002; Mumtaz and Durkin, 1992; U.S. EPA, 2000); synergism resulting in a 50% higher response than additivity could have the same mixture WOE classification and BINWOE value as a synergistic response 200% higher than dose addition would predict. Addition of BINWOE scores does not differentiate between no interaction data ("?",) and evidence supporting additivity ("=") (Mumtaz *et al.*, 1994). A default interactions uncertainty factor of 10 is used but there is no guidance or scientific basis for this value (Mumtaz and Durkin, 1992; U.S. EPA, 2000). Due to these issues with BINWOE calculations, a qualitative mixture WOE evaluation alongside a hazard index developed using target toxicity doses is generally recommended for most mixtures assessments (Pohl *et al.*, 2003).

The mixture WOE procedure is data intensive. As for all interactions approaches, binary toxicity data are scarce and often deficient for the purposes of determining mixture WOE classifications and BINWOE scores (Durkin *et al.*, 1995). The ATSDR recommends assigning a point of departure for including chemicals in the mixture WOE classification. If a chemical's HQ is less than 0.1 or its cancer risk is smaller than 1×10^{-6} , this component may be dropped from the list prior to assessing all binary pairs for potential interaction (ATSDR, 2001a).

Considerable judgment is called for with both the qualitative and semi-quantitative options. The steps in this procedure are fairly complex (U.S. EPA, 2000). Mixture WOE classifications are generally developed by a small group of professionals; classifications must be peer reviewed by a panel or working group (Mumtaz and Durkin, 1992). Since the mixture WOE classifications are works of judgment, numerical representations of these decisions, the BINWOEs, are also judgmental. The BINWOE scores have no absolute numerical significance and can not be used in statistical analysis (Mumtaz and Durkin, 1992; Mumtaz *et al.*, 1994). Professional judgment is required to determine the effect of a qualitative mixture WOE on the mixture HI (ATSDR, 2001a). Both mixture WOE and BINWOE are supposed to be supported by

complete narrations supporting the reasoning used in their development (Johnson and DeRosa, 1995; Mumtaz and Durkin, 1992).

Table 9. Advantages and Disadvantages of the Mixture WOE Procedure

Advantages	Disadvantages
• Incorporates interaction data systematically • Qualitative or semi-quantitative • Non-cancer or carcinogenic endpoints • Consistent with toxicological tests of mixtures • Consistent application between scientists	• Data often insufficient to evaluate binary interactions • BINWOE scores oversimplify chemical interactions • Data intensive • Professional judgment intensive • Requires peer review

Interaction Hazard Index Procedure

The biggest disadvantages of the mixture WOE procedure are considered to be the fact that the mixture HI is multiplied by a single UF (Equation 5) and that the normalized WOE used to modify the UF does not consider the magnitude of the difference from additivity. Based on these concerns, the semi-quantitative mixture WOE equation was modified to the interaction HI procedure (ATSDR, 2001a; Hertzberg and MacDonell, 2002; U.S. EPA, 2000). The mixture WOE assumptions remain; binary interactions are expected to account for the majority of effects in the mixture and the toxicity of the mixture can be described by deviations from dose addition (Hertzberg and MacDonell, 2002). In addition, the interaction HI procedure assumes that the interaction between two chemicals is maximal when the chemicals are present in equally toxic amounts and that as the magnitude of the interaction decreases, the outcome will approach additivity (U.S. EPA, 2000).

The interaction HI incorporates new factors along with these assumptions. The magnitude (M) factor describes the maximum influence one chemical has on the toxicity of another. As seen with the mixture WOE procedure, interactions are not necessarily

“symmetric”; the effect of chemical y on chemical x (M_{xy}) is not necessarily equal to M_{yx} . Real data on the change in toxicity of one chemical in the presence of another is preferred, when available (U.S. EPA, 2000). When direct measurements of toxicity are not found, changes in kinetics or bioavailability can be substituted (Krishnan *et al.*, 1994). The default value for M has been set at 5; studies of binary action have shown the maximum M to be ± 5 . The direction of interaction (synergism or antagonism) is not factored into M itself, but into the weight of evidence score (U.S. EPA, 2000).

The weight of evidence score (B) reflects the strength of the data available on interactions between chemicals. Each interaction (e.g., the strength of evidence that y will influence x or B_{xy}) is assigned a category based on the quality of data indicating that an interaction does/could occur and the relevance of the interaction to human health (Table 10). The direction of interaction then indicates the score given for the B value. Synergistic and antagonistic interactions having the best category of data, unequivocal direction of interaction and being relevant to humans receive a +1 or -1, respectively. For chemical pairs with only animal data or studies of lower quality/relevance, synergistic relationships receive higher relative values of B than antagonistic interactions, reflecting the protectiveness built in to the interaction HI procedure. Again, weight of evidence scores do not have to be “symmetric”; B_{xy} often does not equal B_{yx} (Hertzberg and MacDonell, 2002; U.S. EPA, 2000).

Table 10. Interaction HI Weight of Evidence Scores (B)*

Category	Description	Greater than Additive Score	Less than Additive Score
I	Interaction directly relevant to humans and direction unequivocal	1.0	-1.0
II	Interaction demonstrated in relevant animal studies; Relevance to humans likely	0.75	-0.5
III	Interaction plausible but evidence and relevance to humans is weak	0.5	0.0
IV	Additivity demonstrated or accepted because of weak data	0.0	0.0

*Adapted from Hertzberg and MacDonell (2002) and U.S. EPA (2000)

The symbol theta represents the proportion factor. This factor mathematically incorporates the assumption that interaction will be greatest when chemicals are present in equally toxic amounts. This is accomplished by taking the ratio of the geometric mean of the chemical HQs to the arithmetic mean of the HQs:

$$\theta_{xy} = \frac{(HQ_x \times HQ_y)^{0.5}}{(HQ_x + HQ_y) \times 0.5} \quad \text{Equation 6}$$

θ will approach 1 when the HQs for chemicals x and y are similar and will near 0 when one chemical is present in very toxic amounts compared to the other (U.S. EPA, 2000). These factors are combined together to form the interaction HI (HI_I):

$$HI_I = \sum_{x=1}^n (HQ_x \times \sum_{y \neq x}^n f_{xy} M_{xy}^{B_{xy} \theta_{xy}}) \quad \text{Equation 7}$$

where the component f in the interaction HI is the exposure factor. This factor represents the toxic hazard of chemical y relative to the total hazard of the mixture interacting with chemical x. The exposure factor normalizes the modified magnitude value so that, in the case of additivity or no data, the sum of HQs would equal the mixtures HI (HI_{ADD}) (Teuschler *et al.*, 2000; U.S. EPA, 2000):

$$f_{xy} = \frac{HQ_y}{(HI_{ADD} - HQ_x)}$$

Equation 8

One advantage of the interaction HI procedure is that the interaction effects on each chemical's toxicity are applied to each chemical's HQ prior to the HQs being summed (Hertzberg and MacDonell, 2002). The procedure offers greater use of quantitative data on binary interaction. Magnitude of interaction is incorporated and changing proportions of chemicals in the mixture can be accommodated (ATSDR, 2001a). The effect of chemical x on y may be synergistic while y on x may be antagonistic; "asymmetric" interactions are applied individually, giving each the weight they deserve, instead of canceling out, as they would in the quantitative or semi-quantitative application of the mixture WOE method. More types of chemicals with somewhat different modes of action can be represented; inducers of an enzyme can be incorporated into a target organ interaction HI along with chemicals that are metabolized by the enzyme (U.S. EPA, 2000).

As with all HI procedures, an inherent disadvantage is that the HI is an estimated unitless ratio on which no statistical evaluations or tests for accuracy can be performed (Hertzberg and MacDonell, 2002). Another disadvantage is the shortage of data necessary to calculate an interaction HI. The procedure requires more judgment on the part of the risk assessor, in assigning not only the weight of evidence (B) factor, but also the interaction factor, and requires more data, which are rarely available. Magnitude of interaction is frequently not mentioned or not quantified in published studies (ATSDR, 2001a; Hertzberg and MacDonell, 2002; Mumtaz *et al.*, 1994). The procedure has not been tested for consistency of application between scientists (ATSDR, 2001a). It is, however, undergoing verification through laboratory studies (U.S. EPA, 2004b).

Table 11. Advantages and Disadvantages of the Interaction HI Procedure

Advantages	Disadvantages
• Applies interaction information to each chemical's HQ • Semi-quantitative • Non-cancer or carcinogenic endpoints • Incorporates magnitude of interaction • Accommodates changing proportions of chemicals in mixture • Can incorporate more than one related mode of action	• Data often insufficient to evaluate binary interactions • Data intensive • Professional judgment intensive • HI is an estimated ratio • Consistency of application untested

Pharmacokinetic/Pharmacodynamic Modeling

Physiologically based pharmacokinetic (PBPK) and pharmacodynamic (PBD) models are considered by the EPA as "the most desirable approaches for quantifying toxic effects"; EPA actively encourages their development (U.S. EPA, 2002c). The ATSDR endorses use of PBPK models whenever applicable and is using a PBPK model as the basis for its "Interaction Profile" for benzene, toluene, ethylbenzene and xylene (ATSDR, 2001a; Pohl *et al.*, 1999; 2003). Unfortunately, these models are generally developed to fulfill specific research needs; there are relatively few mixtures models compared to the number of mixtures in the environment (Seed *et al.*, 1995).

PBPK and PBD models can be defined as series of mathematical equations meant to simplify biological and chemical kinetic processes and to describe these processes in a biologically relevant way that allows the modeler to better understand the observed toxicological effects (Robinson and MacDonell, 2004). The description of a human body in a PBPK model must be parsimonious and functional; the model must describe distribution of a chemical or mixture of chemicals between compartments effectively without too cumbersome details. Each model must include key organs or lumped tissue groups, transport between organs through systemic circulation and the transfer of the chemical(s) of concern between the blood and the

tissue through either partitioning or transport mechanisms (Verhaar *et al.*, 1997). The kinetics of the chemicals within each tissue in the model is described by physiological parameters including volume, blood flow and partition coefficients, as well as metabolic capacity if applicable (El-Masri *et al.*, 1997). Using these parameters, mass balance differential equations describe the rate of change of chemical in each tissue or organ and can be used to predict the chemical concentration at the target site (Haddad and Krishnan, 1998; Verhaar *et al.*, 1997).

Mixture models are often an assembly of single chemical models linked pairwise at the point of interaction (e.g., the hepatic metabolism description) (ATSDR, 2001a). In preliminary models, "no interaction" is assumed as a default and response is predicted from the summed target organ dose (Feron and Groten, 2002; Verhaar *et al.*, 1997). When additivity can not predict the response, one or more interactions must be incorporated. Binary interactions for each possible pair of chemicals are added to the model; modeling of pairwise chemical interactions has been shown to be sufficient to describe the effect of the entire mixture (ATSDR, 2001a; Krishnan *et al.*, 2002; Tardif *et al.*, 1997). Interactions may be PK or PD in nature or a mixture may contain chemicals resulting in both types of relationships.

PBPK models describe the absorption, distribution, metabolism and excretion of chemicals and are used to predict the dose of chemical at the site of action (Robinson and MacDonell, 2004; Thomas *et al.*, 2002). Changes in kinetics are the most prevalent interactions observed (Hertzberg and MacDonell, 2002; Robinson and MacDonell, 2004). Kinetic interactions between chemicals can occur during any portion of the chemicals' route through the body (U.S. EPA, 2000); modelers attempt to describe these interactions mathematically. Competitive binding to plasma protein molecules can be modeled from the free concentrations of the competing chemicals and the number of total blood binding proteins (Haddad and Krishnan, 1998). Interactions frequently occur as competition for an enzyme. PBPK models account for binding to a substrate. When competitive chemicals are present, the inhibition constant can be determined experimentally from pairwise interactions (Tardiff *et al.*, 1997) or it

can be assumed equal to the interacting chemical's affinity constant (Andersen *et al.*, 1987; Dennison *et al.*, 2004; El-Masri *et al.*, 1997; Krishnan *et al.*, 2002). Metabolism of one chemical may deplete the amount of enzyme or co-factor available for metabolism of a second chemical. Michaelis-Menten kinetics are used to describe the interaction of chemicals competing for the same enzymes and can determine the concentrations of chemicals leading to saturation of the pathway (Casseo *et al.*, 1998; U.S. EPA, 2000). It is important to note that the tissue response to a chemical is not expected to change in a PK interaction; in a PBPK model, the dose of chemical delivered to the tissue of concern will change due to an interactive mixtures exposure (Haddad and Krishnan, 1998; Andersen and Dennison, 2004).

Interactions at the PD level, however, change tissue response to a given tissue dose (Haddad and Krishnan, 1998; Andersen and Dennison, 2004). PBPD models focus on biological effects at the tissue of concern and attempt to describe tissue response in a quantitative way (El-Masri *et al.*, 1997; Thomas *et al.*, 2002). Responses often include changes in gene expression, resulting in adaptive responses such as up-regulation of enzymes (e.g., P450), cell proliferation or even cell death (Robinson and MacDonell, 2004; Thomas *et al.*, 2002). PD interactions may occur at the same mechanism or by different related mechanisms; one chemical may elicit cell death while a second impedes cell regeneration or one carcinogenic chemical may be an initiator while another acts as a promoter (El Masri *et al.*, 1997; Seed *et al.*, 1995).

With physiologically based models, extrapolation is "their greatest strength" (Connolly, 2001). PBPK models can be used to extrapolate between high doses seen in experiments and lower doses more commonly seen in human environmental exposures. Route to route extrapolation is also possible for a well constructed PBPK model (Connolly, 2001; Haddad and Krishnan, 1998; Health Council of the Netherlands, 2002; Henschler *et al.*, 1996). PBPK models can be used to examine time frame overlaps such as subsequent exposures of slowly

metabolized/persistent chemicals or the formation of toxic metabolites with the parent compound still in the body (ILSI, 1999; U.S. EPA, 2000).

Prediction is another strength of PBPK and PBPD models (Table 12). In a single chemical assessment, an animal model is used to estimate a target tissue dose associated with a specified level of adverse effect (similar to a BMD); a companion human model back-calculates the exposure level necessary to cause the same response in humans (Haddad and Krishnan, 1998; U.S. EPA, 2000; Yang *et al.*, 1998). In a mixture assessment, parallel human and animal models can be used in the same way to predict the exposure level of a defined mixture that would result in a human interaction threshold. An interaction threshold is defined as the point on a mixture dose-response curve where additive effects transition to non-additive (synergistic or antagonistic) results. Human exposure below this threshold would result in purely additive effects of the mixture components (El-Masri *et al.*, 2004; Yang *et al.*, 1998). Uncertainty in animal to human extrapolation can be greatly reduced by using validated PBPK models in this fashion (Krishnan *et al.*, 1994; Sexton *et al.*, 1995). Sources of uncertainty and variability can be identified through the use of sensitivity analysis (ILSI, 1999).

PBPK models may be utilized in risk assessment through the calculation of PK-HI ratios for a mixture. This ratio is the sum of PK-HQs, which are the ratio of target tissue dose of a component during mixture exposure divided by the target tissue dose of that chemical if exposure had occurred at singly the RfD. The PK-HI ratio for a mixture could be the same as a traditional additive HI (HI_{ADD}) if no interactions occur; however, the ratio would be less than HI_{ADD} in the case of antagonistic interactions and greater than HI_{ADD} in the case of synergism between components (Haddad *et al.*, 2001; Liao *et al.*, 2002).

Models can be used to predict concentrations for mixture components that could result in interactive effects and the likely magnitude of the interaction. These predictive exercises could help determine dose levels for future toxicological studies. The use of PBPK models in experimental planning could result in more cost-effective investigation using fewer laboratory

animals (Dobrev *et al.*, 2002; NORA, 2004; Andersen and Dennison, 2004). Once developed, a PBPK or PD model evaluates changes in doses or exposure scenarios quickly and cheaply on a desktop computer (Connolly, 2001). Models are often most useful when developed in conjunction with laboratory studies (Teuschler *et al.*, 2002). Joint development of models and studies can lead to better understanding of the mechanism, especially complicated behaviors such as suicide inhibition (i.e., inhibition of self-metabolism along with metabolism of other chemicals) (Andersen and Dennison, 2004; Fisher *et al.*, 2004).

The major disadvantage of PBPK and PBPD models is that they are data intensive. Exposure data requirements can include permeability coefficients for dermal exposure and particulate size or measurement of the volatile portion for inhalation exposures. A realistic ventilation rate must be determined for an inhalation model. Literature values of blood flow and tissue volume must be located for each compartment (Krishnan *et al.*, 1994; NORA, 2004; U.S. EPA, 2000; Verhaar *et al.*, 1997). Some physiological data are species specific but not changed by different chemicals and so do not have to be newly determined for each model; however, chemical specific parameters have to be experimentally derived or calculated for each new component. Partition coefficients can be found in publications, quickly evaluated using *in vitro* methods or be calculated in a number of ways with varying degrees of success depending on chemical properties and method used (Sterner *et al.*, 2004). Metabolic capacity, metabolic rates and protein binding constants can also be found in literature, determined through *in vitro* or *in vivo* tests, or be calculated using quantitative structure activity relationship (QSAR) models (Cassee *et al.*, 1998; Krishnan *et al.*, 1994; NORA, 2004; Verhaar *et al.*, 1997). To make data gathering even more complex, chemical interactions can affect tissue:blood partitioning (carbamates increase lead uptake into the brain), ventilation rate (decreased by central nervous system depressants) or metabolic capacity (increased by enzyme induction) (Dennison *et al.*, 2004; Krishnan *et al.*, 1994). Experimental studies having time course data can be found in the literature but often studies must be designed specifically for kinetics research (e.g., gas

chamber uptake studies for volatiles) to be of best use in model development (Andersen *et al.*, 1987; Dennison *et al.*, 2003). Binary interaction study data are invaluable to mixture model development (Tardif *et al.*, 1997). Finally, prior to peer and regulatory acceptance, models must be validated. Data from experimental studies not used during development of the model must compare favorably with model predictions (ATSDR, 2001a; Sexton *et al.*, 1995). While kinetics data for individual components and binary pairs can be used for model development; toxicity data from the mixture should be used as part of the validation process (El-Masri *et al.*, 2004). Following completion, models must be evaluated before being used in a risk assessment; appropriate documentation throughout the model development process is necessary for this evaluation and peer acceptance (Clark *et al.*, 2004). Because models are highly data, resource and knowledge intensive, modeling is limited to mixtures representing sufficient human health risk to offset the cost of studies and people hours necessary for their development and validation (Hertzberg and MacDonell, 2002; ILSI, 1999; Robinson and MacDonell, 2004).

The more mixture components, the greater the data demand. Lack of data can make complex models nearly impossible to program. In addition, the more parameters and equations in a mixture model, the greater the propagated error from rounding. Higher precision levels have to be used to keep model predictions as accurate as possible. If the model becomes too complex, a degree of precision may be needed that can not be provided by the software package (Verhaar *et al.*, 1997). For these reasons, a process called data lumping is used for complex mixtures. This approach has been borrowed from the petroleum refining industry and subsequently used in petroleum product PBPK models. Similar components are lumped together based on chemical (e.g., partition coefficients) or biological (e.g., expected mode of action) characteristics. The lump is then treated as a single chemical and target tissue levels of the lumped set can be predicted by a single chemical model format. A lump will not behave chemically like a simple component; QSAR programs may be used to estimate average or representative partition coefficients or a molecular weight for a lumped component set (ATSDR,

2001a; NORA, 2004; Verhaar *et al.*, 1997; Yang *et al.*, 1998). If a single or a few components are of interest to the modeler, they may be subtracted from the main lump and modeled individually, interacting with the lumped portion and each other on a binary level. However, caution must be used as the composition and physical characteristics of the lump may change over time as some components are metabolized or volatilized (i.e., exhaled) (Dennison *et al.*, 2003; 2004).

A disadvantage common to all PBPK and PBPD models is the fact that some parameters have to be estimated and assumptions must be made. The process of fitting data to estimate parameters is not standardized; generally simulation curves are fitted to data by "eyeball" (Hertzberg and MacDonell, 2002). Assumptions can get the modeler in trouble if one forgets what portions of the model are data-based and what are speculation-based. Predictions made by the model are only as good as the mathematical descriptions of mechanisms and interactions are accurate (Connolly, 2001).

Table 12. Advantages and Disadvantages of PBPK/PBPD Modeling

Advantages	Disadvantages
• Applies binary interaction information • Quantitative • Non-cancer or carcinogenic endpoints • Incorporates magnitude of interaction • Accommodates changing proportions of chemicals in mixture • Can incorporate more than one related mode of action • Extrapolates between doses, dosing schemes, routes of exposure • Predicts tissue dose • Facilitates and reduces uncertainty in animal to human extrapolation • Used to calculate PK-HI values for the mixture	• Data often insufficient to evaluate binary interactions • Data intensive • Documentation intensive • Often requires specialized studies and assays • Costly • Time consuming • Estimations and assumptions inherent to biologically based modeling

TOXICOLOGY OF CHEMICAL MIXTURES RISK ASSESSMENT FOR ECOLOGICAL RECEPTORS

Incorporation of Mixtures in ERA

Mixtures assessments are expected to be incorporated in every ecological risk assessment (U.S. EPA, 1997a; Wentsel *et al.*, 1994; 1996). This is clear by the DoD definition of ERA as the "qualitative or quantitative appraisal of the actual or potential impacts of stressors on plants and animals at a site" (Simini *et al.*, 2000). Stressors are chemical, biological or physical entities that can result in adverse effects; in quantitative assessment, chemicals are generally the only stressors considered. The functions of ERA are threefold: characterize the current or potential risk to ecological populations at the site, identify the contaminants that are associated with unacceptable risk and produce data useful for determining cleanup options (U.S. EPA, 1997a).

The U.S. EPA has proposed an eight step process for performing ERAs (U.S. EPA, 1997a). The DoD uses a tiered approach that incorporates many of the same steps, but not in such a formal order. The two approaches are compared in Table 13. The Army and Air Force completely share their tiered ERA process. Tier 1 is a screening level risk assessment similar to Steps 1 and 2 of the EPA process. Tier 2 is a formal baseline risk assessment while Tier 3 reiterates Tier 2 with more site-specific data (Simini *et al.*, 2000). The assessment may be halted after completion of any tier depending on decisions by the regulators and site management regarding the benefit of more site-specific information potentially gathered in the following tier (Wireman, personal communication, 2004). The Navy follows Tiers 1 and 2; their Tier 3 consists of evaluation of remedial alternatives to reduce ecological and human risk at the site to acceptable levels (Simini *et al.*, 2000). For ease of reference, Army and Air Force tiers are used in this review.

Table 13. Comparison of U.S. EPA Steps for ERA and DoD Tiered ERA

U.S. EPA (1997a)	DoD (Simini <i>et al.</i> , 2000)	
Step 1: Screening-level problem formulation and ecological effects evaluation	Tier 1 Screening Level Risk Assessment	
Step 2: Screening-level exposure estimate and risk calculation		
Step 3: Baseline risk assessment problem formulation	Tier 2 Baseline Risk Assessment	Tier 3 Site-specific Risk Assessment (Army and Air Force)
Step 4: Study design and data quality objective process		
Step 5: Field verification of sampling design		
Step 6: Site investigation and analysis phase		
Step 7: Risk characterization	Risk Characterization	
Step 8: Risk management	Risk Management	

Mixtures risk assessment is incorporated into every tier of the ERA process. The Tier 1 screening level assessment is a conservative, paper-based estimation of site ecological risks. This estimation is biased toward conservatism; chances that a risk is overlooked when there is indeed an ecological impact should be minimized. Historical contamination at the site, previous sampling results and literature studies of contaminant toxicity are used to judge potential hazards at the site. Conservative fate and effects models are used to predict potential exposure levels to site receptors and bioavailability estimations used to determine possible food chain effects. Hazard indices, discussed later in this review, are calculated to help decide whether a baseline risk assessment should be performed (Simini *et al.*, 2000; U.S. EPA, 1997a). These HIs are generally based on maximal contaminant levels or 95th upper confidence level of the arithmetic mean (Fairbrother, 2003; Wireman *et al.*, 2003). Risk managers and regulators have to decide whether to proceed to a Tier 2 evaluation based on the screening level HIs and the confidence that the HIs truly represent potential risk at the site (Wireman, personal communication, 2004).

A Tier 2 or baseline risk assessment is initiated due to excess risk posed by the screening level HIs or the need to reduce uncertainty in the assessment. Tier 2 should include

a comprehensive literature assessment of contaminant toxicity and mechanisms, site media sampling and incorporation of these data into fate and effects models, and sampling of site food items to incorporate into bioavailability estimations. As site data are substituted for conservative assumptions made in the screening level RA, the baseline RA becomes more realistic (Fairbrother, 2003; Simini *et al.*, 2000; U.S. EPA, 1997a). Many site specific refinements in the baseline risk assessment focus on the exposure assessment. Biomarkers may be assayed to confirm receptor exposure and determine if exposures are similar to those predicted using fate and effects models. The DoD is developing a Terrestrial Biomarker Database to assist in this confirmation (Wireman *et al.*, 2003). Default models can be upgraded to more sophisticated fate, transport and persistence models to suit site conditions (Cahill *et al.*, 2003). Although the focus of Tier 2 is generally on reducing uncertainty in exposure, significant uncertainty remains in the toxicity assessment (Fairbrother, 2003). Short-term whole mixtures toxicity tests, discussed later in this review, may help fill toxicity data gaps and confirm ecological exposure and effects (Simini *et al.*, 2000; Wentsel *et al.*, 1994; 1996; Wireman *et al.*, 2003).

Tier 3, an even more site-specific risk assessment, reiterates the assessment in Tier 2 but more site specific data are collected to reduce uncertainty. Longer term (i.e., 6 months or more) whole mixtures toxicity tests and more complex (i.e., population and ecosystem level) field studies may be utilized to fill data gaps (Simini *et al.*, 2000; Wireman *et al.*, 2003). Tier 3 will often focus on less polluted areas to determine if they represent ecosystem risk; portions of the site that are highly polluted (i.e., hot spots) can be assessed adequately in Tier 2 or even Tier 1 (Wentsel *et al.*, 1994; 1996).

Following completion of Tier 2 or Tier 3 studies, the risk characterization process begins. Like HHRA (U.S. EPA, 1989), the risk characterization step includes the calculation of risk levels (outside of the screening level assessment calculations, which are incorporated into Tier 1) (U.S. EPA, 1997a). A weight of evidence process is used to link measurement endpoints to assessment endpoints to evaluate whether a significant ecological impairment has occurred or

may occur. This process should be quantitative to the extent possible, but a qualitative assessment is acceptable. Weighting or ranking of measurement endpoints is one way to make this process semi-quantitative (Burton *et al.*, 2002).

Assessment endpoints identify the specific components of the site ecology that are to be protected and what that protection entails (e.g., reduction of key species population, destruction of a specific ecosystem). Assessment endpoints are chosen so that they will aid in the decision making process by defining ecological impairment. Measurement endpoints are measurable responses to stressors (decreased survivability of test species A, reduced reproduction in species B). Measurement endpoints relate to an assessment endpoint and can be used to infer the likely outcome of the assessment endpoint (Simini *et al.*, 2000; U.S. EPA, 1997a). The importance of each endpoint should be decided among the site stakeholders prior to the beginning of any studies. Measurement endpoint weight can depend on data quality, intensity of toxic response and concurrence between related measurement endpoints. Evidence of chemical interaction influencing the toxicological outcome of a given measurement endpoint could be incorporated into the weight of that endpoint. The purpose of any weight of evidence process is to give the risk manager an overall idea of the risks to all portions of the ecosystem and an synopsis of the uncertainties involved in estimating those risks (Simini *et al.*, 2000).

ERAs can involve entire ecosystems or watersheds (U.S. EPA, 1997a) and thousands of species can be affected. Interaction occurrence, direction and magnitude can differ between species (Hertzberg and MacDonell, 2002; Seed *et al.*, 1995). Factor in more than 100 identified chemicals at a typical U.S. EPA hazardous waste site (Johnson and DeRosa, 1995) with the paucity of interactions data and the result is too many variables to allow assessment of potential binary interactions. However, some consideration has been given to assessing interactions data between chemical classes. Components of one class may be assumed to interact with constituents of another class in a similar fashion. Class definitions would be judgmental; they could include structurally similar compounds such as ketones or be made up of chemicals with

the same mode of action (Ah receptor mediated toxicity). Interaction studies could be compiled cross-species to circumvent the data deficiency problem. Although this method would be qualitative at best, it could provide a systematic evaluation of potential interactions in an ERA (Deneer, 2000; Durkin *et al.*, 1995).

Like HHRA, there are two types of approaches for assessing the toxicity of mixtures in ERA. The component or bottom-up approach is preferred for Tier 1 screening level assessments. Top-down or whole mixture methods are generally used to generate site specific data in Tiers 2 and 3 (Fairbrother, 2003).

Risk Calculation Procedures for Component Dose Addition Approaches

Component approaches are generally used in a Tier 1 assessment to relate concentrations of site chemicals to toxic doses in an attempt to predict biological effects. An advantage of using component approaches for Tier 1 is that existing data may be used (Ankley and Mount, 1996; Fairbrother, 2003). Tier 1 assessments are largely carried out in the office (Simini *et al.*, 2000). Field studies, toxicity studies and biomarker gathering can be expensive. Component approaches generally rely on toxicity data from the literature and past site sampling results. Also, fate of individual chemicals is better understood and can be modeled easily to project exposure concentrations for Tier 1 assessments. In addition, component approaches are quantitative and can be used to establish remediation goals for specific contaminants (Ankley and Mount, 1996; Smolders *et al.*, 2003; Van Leeuwen *et al.*, 1996).

The main disadvantage of component approaches is that not all toxicants present at the site are measurable. Contaminants missed by chemical screening analyses or below the detection limit can contribute to toxicity (Ankley and Mount, 1996; Smolders *et al.*, 2003; Van Leeuwen *et al.*, 1996). A component approach (additive HI) was shown to underestimate the toxicity of a lead smelter effluent due to metals or other toxicants in the mixture that were below detection limits; this underestimation was not due to interaction as an artificial mixture

containing the measured metals produced the predicted additive toxicity (Ross and Bidwell, 2003).

Further disadvantages of component approaches include the assumptions of "no interaction" and bioavailability. Bioavailability is frequently ignored when component approaches are used (assumed 100% bioavailable), especially in screening level assessments (Smolders *et al.*, 2003; Van Leeuwen *et al.*, 1996). A default number, such as 10% for metals, may be assumed (Schoof, 2003). Chemical-chemical interactions are assumed to be negligible. Toxicity data may not be available for site contaminants or data may be insufficient to address the number of species necessary to characterize site risk. Complete toxicity testing batteries are expensive. Further, predicted toxicity from a component approach may or may not reflect biological impairment at the site (Smolders *et al.*, 2003; Van Leeuwen *et al.*, 1996).

A similarity to HHRA is the usefulness of stressor-response curves in ERA. The shape of the dose-response curve indicates if there is an effective threshold and what a "no interaction" response should be. For the main dose addition assumption to be upheld, chemicals are required to have similarly shaped curves. Sufficient dose-response data to build a curve is needed for input into effects models to predict impact (Niederlehner *et al.*, 1998; U.S. EPA, 1998). Unfortunately, ecological toxicity and mode of action data on a single chemical are rarely available in sufficient quantity in a single species to produce a stressor-response curve, let alone the multiple curves for various site chemicals and species types needed to assess site risk (Grimme *et al.*, 1996; Smolders *et al.*, 2003; Van Leeuwen *et al.*, 1996; Vighi and Calamari, 1996).

Dose addition and response addition are defined the same in ERA as in HHRA. When chemicals behave through different modes of action, they can be treated independently using response addition (ECETOC, 2001; Escher and Hermens, 2002). Mixtures of 14 to 16 chemicals having independent modes of action were found to be response additive in algae and bacteria in acute growth and bioavailability assays. Dose addition overestimated the risk (Faust

et al., 2000). However, response addition does not work in all situations of independent action. Frequently, the "response" summed is mortality (ECETOC, 2001). The response addition approach may not hold when the endpoints are total body effects (e.g., lethality), not the subtle cellular changes used to mark responses in HHRA (e.g., a mutation event). Response addition has been shown to consistently underpredict the toxicity of some mixtures of dissimilar chemicals in acute algae survivability assays while dose addition over- or underpredicted toxicity within an accepted factor of 2.0 (Grimme *et al.*, 1996). Because response addition does not work dependably in ERA and as dose addition provides good approximation of risk regardless of mode of action, response addition is not frequently used in ERA. Dose addition is the only category of component approach discussed in this review.

Dose addition is the most commonly used component method for predicting mixture toxicity (Niederlehner *et al.*, 1998). Toxicity of related chemicals or those having the same mode of action can be summed using additivity or dose addition (ECETOC, 2001; Escher and Hermens, 2002). Additivity has been confirmed in various species for PAHs, triazine congeners, BTEX (benzene, toluene, ethylbenzene and xylenes) and many metals (Ankley and Mount, 1996; Dyer *et al.*, 2000; Faust *et al.*, 2000; Niederlehner *et al.*, 1998; Ross and Bidwell, 2003; Vighi and Calamari, 1996). Additivity occurs even when components are present below the individual chemical lowest observable effect concentrations (LOECs). Concentrations can be as low as 0.25% of an EC₅₀ (effective concentration for 50% of the exposed organisms) (Escher and Hermens, 2002; Niederlehner *et al.*, 1998; Vighi and Calamari, 1996). Additivity was used to explain cumulative metals toxicity to striped bass in Chesapeake Bay; Al, Cd, Cr, Cu and Zn were present in the Bay below regulatory levels but dose additive predictions of toxicity correlated well with observed effects (Logan and Wilson, 1995).

Dose addition is accepted as the approach to be used for baseline toxicity of chemicals in aquatic environments (Deneer, 2000). Also known as narcosis or non-specific activity, baseline toxicity results from the partitioning of organics into membranes and their adsorption

into macromolecules. Disturbance of membrane integrity and function leads to morbidity or mortality. Toxicity is related to the membrane concentration of the chemical; therefore, a chemical's octanol:water partition coefficient, which represents membrane solubility, is correlated with the mortality EC_{50} . The effects are reversible, provided the membrane concentration fails to reach the internal effect threshold; this threshold is specific to species, size (surface:volume ratio), lipid content and other factors (Altenburger *et al.*, 2003; Escher and Hermens, 2002). Most organic compounds have specific toxicity at higher concentrations but low levels can result in non-specific effects (Dyer *et al.*, 2000; Niederlehner *et al.*, 1998). Because distribution, partitioning and adsorption occur at all concentrations, small concentrations of a large number of organics can result in baseline toxicity (Altenburger *et al.*, 2003). Dose addition of mixtures containing up to 50 nonspecific toxicants has been confirmed in multiple aquatic species, providing mode of action additive concentrations were below thresholds for specific effects (Escher and Hermens, 2002). It is estimated that the majority of chemicals on the Toxic Substances Control Act inventory and most industrial chemicals (e.g., alcohols, ketones, ethers, aromatic and aliphatic hydrocarbons) have narcotic effects to aquatic life (Logan and Wilson, 1995; Van Leeuwen *et al.*, 1996).

Frequently, additivity is found to be reasonably applicable to various mixtures, even when constructed of theoretically independent chemicals (Altenburger *et al.*, 2003; Deneer, 2000). This is particularly true for acute toxicity of a large number of components at low concentrations or chronic toxicity of compounds in the environment below their individual LOECs (ECETOC, 2001; Escher and Hermens, 2002). For 20 of 24 mixtures containing similar and dissimilar chemicals, dose addition predicted toxicity within 95% confidence limits of actual toxicity observed (Niederlehner *et al.*, 1998). A review of aquatic toxicity of 202 binary insecticide, herbicide and fungicide mixtures found that 92% were within the bounds of additivity (i.e., between 0.5 and 2.0 toxic equivalents when the expected toxicity from dose addition was

1.0). Of the mixtures considered to be independent, 90% were within these additivity limits (Deneer, 2000).

Interactions between chemicals do occur. However, there are few examples in the literature of effects greater than three times the effects expected from additivity (ECETOC, 2001). Non-additive interactions increase as dose of the mixture increases but may be significant at low, environmentally relevant doses (Haas *et al.*, 1996). For regulatory purposes, dose addition is considered to provide the reasonable worst case estimate of toxicity for similar and dissimilar mixtures (Faust *et al.*, 2000; Grimme *et al.*, 1996). In order to move beyond worst case scenarios, dose addition must be done by mode of action for specific toxicity. There is a lack of consensus on how to group chemicals for additive toxicity estimation. Due to the lack of mechanistic data for ecological toxicants, target organ groupings are most frequent but mode of action or even mechanistic groupings would be more accurate. QSAR models may be used to hypothesize on the mode of action or potential site of toxicity (Grimme *et al.*, 1996; Vighi and Calamari, 1996).

Hazard Index Procedure

As in HHRA, the hazard index procedure is the most commonly used component approach in ERA. Hazard indices are the summation of hazard quotients, also known as toxic units, which relate known or predicted exposure concentrations to some measure of toxicity. The metric of toxicity often changes with the purpose of the assessment. HQs constructed from toxicity values for species from the same receptor taxon are considered comparable. Chemicals should be assigned to mode of action or target organ groups and HIs should be summed for each group. A HI value greater than 1.0 is generally assumed to indicate potential risk at the site. A HI less than 1.0 does not necessarily indicate an absence of risk, depending on the severity of effect used to calculate HQs; if an acute effect level is the common denominator in the HQs and a HI less than 1.0 is calculated, this does not indicate that chronic

effects are unlikely to occur (Ankley and Mount, 1996; Simini *et al.*, 2000; U.S. EPA, 1998; Wentsel *et al.*, 1994; 1996).

The HI procedure is primarily used in Tier 1 and 2 assessments (Wentsel *et al.*, 1994; 1996). In screening level assessments, emphasis is often placed on use of available data to quickly estimate potential risk at the site. The European Commission recommends the use of an HI for aquatic toxicity assessment of new and existing chemicals. Predicted exposure concentrations are divided by predicted no effect concentrations (PNEC); PNEC are calculated by dividing the lowest lethal EC_{50} for the species by 1000 (Villa *et al.*, 2003). In the U.S., the exposure level may be calculated from the highest measured concentration at the site for a Tier 1 assessment. The health effect level may be an expedient value such as the Ambient Water Quality Criterion (Simini *et al.*, 2000) or, following a literature search and review, a LOAEL divided by a UF of 10 or a NOAEL can be used. A NOAEL is preferred to ensure that ecological risk is not underestimated; similarly, chronic health values are preferred over acute test values and studies using a similar route of administration to the environmental route of exposure are preferred over dissimilar routes (U.S. EPA, 1997a).

USACHPPM has developed non-site-specific terrestrial wildlife toxicity reference values (TRVs) for use as the denominator in HQ calculations for screening level evaluations. TRVs are similar to human RfDs and are meant to signify an individual toxicity level (i.e., measurement endpoint) that should represent a population effect (i.e., assessment endpoint). Following a comprehensive literature search, a written toxicity profile and documentation of literature values leads to the development of TRVs for mammalian, avian, amphibian and reptilian receptors, if sufficient data are available. Data requirements include a minimum of two chronic NOAELs and two LOAELs from three high quality studies per taxonomic order. Two taxonomic orders per receptor class (e.g., mammal) must be represented. Exposure routes must match between studies and be valid to site exposure scenarios. If sufficient data are unavailable, an uncertainty factor approximation approach is substituted, providing a NOAEL and LOAEL are available from

a relevant study. UFs range from 1 for a chronic NOAEL-based TRV to 100 for a LD₅₀-based TRV. Approximation TRVs are assigned qualitative confidence levels (low, medium or high) but do have the quantitative confidence assigned to data-derived TRVs (USACHPPM, 2000).

Data-derived TRVs are never single numbers but are ranges designed to stretch from no observed response to low toxic responses in the population. Either the LOAEL-NOAEL process (i.e., most sensitive LOAEL and NOAEL values for the receptor class) or the BMD process (i.e., the modeled 10% effective dose (ED₁₀) and the 95% confidence interval on the ED₁₀) may be utilized to estimate the upper and lower values for the data-derived TRV range. Both of these processes are used in HHRA. The BMD process is data intensive but uses the entire dose-response curve for the receptor class. The LOAEL-NOAEL process is less resource intensive but more dependent on good study design. Either the lower or upper TRV value may be used to calculate the HQ; the confidence in the resulting HI is dependent on whether the more or less conservative value is used (USACHPPM, 2000).

In Tier 2, HQs are calculated using better estimates of exposure levels and health effects. Instead of worst case hot spot concentrations, exposure estimates should be more realistic (e.g., 95% confidence limit values) for chronic health effects (Simini *et al.*, 2000; U.S. EPA, 1997a). At some sites, unexpected acute toxicity results from pulsatile releases. Storm event sampling methods may be employed to capture these exposure levels (Dyer *et al.*, 2000). Instead of expedient Ambient Water Quality Values or conservative TRVs, toxicity values should be developed following an extensive literature search and based on NOAELs from the types of receptors present at the site. Toxicity tests can be designed for baseline risk assessments if sufficient data are not available for the various species needed to assess risk at the site. As confidence in the values used in each HQ increase, the resulting HI becomes a more certain indicator of risk at the site (Simini *et al.*, 2000; U.S. EPA, 1997a). Although toxicity testing is expensive and increases the duration of the assessment, the increased confidence in site specific values should far outweigh the cost of studies needed (Fairbrother, 2003).

An advantage of the HI approach is its familiarity and ease of use (Table 14). HI is an inexpensive and efficient procedure. Another advantage of the hazard index is that it uses available literature data whenever applicable (U.S. EPA, 1998). These properties make it the best procedure for assessing risks of mixtures at a screening or baseline level (Wentsel *et al.*, 1994; 1996). During the first stage of a screening assessment, additivity and mode of action assumptions may be ignored and all component HQs can be added together for a receptor species to determine if risk is likely from the pathway in question (Wireman, personal communication, 2004). If excess risk is indicated through this rapid screening, mode of action or target organ HIs should be constructed to determine if risk is still implied (Simini *et al.*, 2000).

The lack of sufficient data is potentially the biggest disadvantage in using the hazard index procedure. Toxicity values for each chemical are needed for different trophic levels (e.g., algae, daphnid, fish). HQ ratios must be comparable to be summed into a HI (Villa *et al.*, 2003). A HI ratio is most defensible when the biological endpoint is the same for all chemicals included (Ankley and Mount, 1996). Frequently, acute EC₅₀ data are more available than chronic no observed effect concentrations (NOECs); known chronic effects are ignored because one or more chemicals included in the HI did not have a comparable chronic toxicity value (Villa *et al.*, 2003). Chronic values may be estimated from acute data using an UF of 10. However this may overestimate risk; dividing EC₅₀ values by 10 can result in a predicted NOEC value lower than the actual NOEC value (van den Brink *et al.*, 2002). Alternatively, QSAR may be used to estimate an endpoint toxicity value for a chemical (Ankley and Mount, 1996). For very poorly studied chemicals where toxicity values exist for a few species, the lowest toxicity value may be divided by an UF ranging from 10 to 1000. This UF is arbitrarily assigned based on the number of taxa tested, the range of toxicity values, the duration of the studies (i.e., acute or chronic) and the conservativeness of the value (i.e., NOECs are more conservative than EC₅₀s) (Fent, 2003).

Effects seen in aquatic receptors are not applicable to other species, scenarios or even to different endpoints in the same species (U.S. EPA, 1998). Much of the chemical toxicity

database for ecological receptors is based on aquatic toxicity tests. Frequently, there are sufficient aquatic data but insufficient sediment toxicity data (Villa *et al.*, 2003). Benthic toxicity values for metals and nonpolar organics may be estimated from aquatic values using the equilibrium partitioning method. This method assumes that the only exposure to sediment dwelling organisms is through pore water; exposure is dependent on the organic carbon content of the sediment and the partitioning of the chemical to organic carbon (MacDonald *et al.*, 2000; Villa *et al.*, 2003). Marine aquatic and sediment organisms differ in sensitivity from freshwater; salinity increases the toxicity of organophosphate insecticides and some metals (Ross and Bidwell, 2003; Villa *et al.*, 2003).

The lack of appropriate data for the HI procedure is a related disadvantage. Wildlife toxicity values are frequently expressed in mg/kg food. In order to convert these values to mg/kg/day, a food consumption value, often not reported for the specific study, is needed. Literature values of food consumption may not reflect actual consumption levels due to growth or palatability issues. Different endpoints and assays used to develop toxicity values can cause order of magnitude differences in resulting HQs (Fairbrother, 2003). If commonly accepted toxicity values such as USACHPPM TRVs are not available then site specific health effect levels are developed; these values tend to vary widely between locations (Wireman *et al.*, 2003).

Frequently, HIs do not incorporate any estimation of bioavailability or bioaccumulation, unless it is incorporated into the model used to predict exposure levels (Fent, 2003). Background metal concentrations are not incorporated either. Metal exposure may not equate with toxicity at the site due to species adaptation to the background. This can result in overprediction of toxicity not correlated with site field assessments (e.g., index of biotic integrity) (Dyer *et al.*, 2000).

The HI procedure results in a single ratio estimate; HIs are not appropriate for use in statistical analysis or probability of effect calculations (Simini *et al.*, 2000). Probabilistic risk assessment allows the quantification of risk through the use of exposure concentration and

single species toxicity data distributions, instead of single values. Probabilistic risk assessment is encouraged in ERA, especially in Tier 2 and 3 assessments when increased site complexity and cleanup costs warrant the extra cost and effort of such an assessment (George *et al.*, 2003; Simini *et al.*, 2000; Wentsel *et al.*, 1994; 1996). When probabilistic assessment is applied to additive mixtures, the HI_{ADD} is used as the probability distribution mean with an assumed lognormal distribution. The variance of this distribution is predicted using the variance values from individual chemical exposure concentrations and the variance from component toxicity values. Risk is then calculated by the normal probability function (Logan and Wilson, 1995). Although probabilistic assessments can be used to help identify the proportion of the population at risk on the site (Wireman *et al.*, 2003) and foregoes the need for arbitrary uncertainty factors, the process is data and computation intensive (Wentsel *et al.*, 1994; 1996).

HIs do not incorporate uncertainty in any quantitative fashion (U.S. EPA, 1998). This uncertainty is frequently augmented by the extrapolation between test species and site receptors; often the extrapolation is based solely on bodyweight without consideration of additional physiological differences (Fairbrother, 2003). Uncertainty in the exposure assessment portion is compounded by sequential use of models to predict exposure concentrations and then to estimate bioaccumulation (Menzie *et al.*, 1992).

Table 14. Advantages and Disadvantages of the Hazard Index Procedure

Advantages	Disadvantages
• Predicts well for similar and dissimilar mixtures • Uses available toxicity data • Familiar and easy to use • In screening assessments, can add all components to determine potential risk • Probabilistic assessment can be used to estimate risk likelihood	• Interactions not considered • Data insufficient to verify dose-response curves and mode of action • Toxicity data often unavailable • Data often inappropriate for assessment endpoint in question • Often excludes bioavailability, bioaccumulation, background metal concentrations • Neglects compounds below detection limit

Toxic Equivalent Factor Procedure

As in HHRA, toxic equivalency factors are assigned for congeners based on their toxicity compared to that of a reference compound. Toxic equivalents for congeners are calculated based on the chemical's TEF and the amount of that component found in the site environment. Toxicity of the congener mixture is estimated by the summation of toxic equivalents (George *et al.*, 2003). The TEF procedure is being used for ERA in several countries (Sanderson and van den Berg, 1999). The World Health Organization recommends the use of TEFs based on rodent toxicity studies for wild mammals (Schroder *et al.*, 2003) and has developed some TEFs for birds and two classes of fish (U.S. EPA, 2000). Studies in fish and wildlife exposed to dioxins, furans and PCBs support the use of additivity and TEFs (U.S. EPA, 2004b).

The assumptions for TEF in ERA remain the same as those in HHRA. The congeners must have the same mode of action, parallel dose response curves, one major toxic pathway and a common structural make-up (George *et al.*, 2003). Doses of congeners must be additive; the dose-response curve for the reference compound should be well characterized (Fent, 2003). The class of compounds must also be persistent and relevant to environmental contamination in order to warrant TEF development (Sanderson and van den Berg, 1999).

Human TEFs have all been based on Ah receptor activity (U.S. EPA, 2000). TEFs used for wildlife toxicity to dioxin, dioxin-like PCBs, furans and PAHs are also based on Ah mediated toxicity. Response to Ah receptor binding differs widely among environmentally relevant species, depending upon the original function of the cell in which Ah activity is induced. A major response to Ah activation in hepatocytes is cytochrome P-450-1a (CYP1a) induction. The metabolic activity of this enzyme can be measured by ethoxyresorufin-O-deethylase (EROD) levels. EROD is a sensitive and rapid response to Ah receptor binding and has been used to determine TEFs in fish species (Sanderson and van den Berg, 1999). Induction equivalency factors (IEFs) have been developed for 19 PAHs, 12 nitrated PAHs and 12 azaarenes using a

fish (*Poeciliopsis lucida*) liver cell line. Toxicity for these congeners was based on their EC₅₀ levels with dibenz[a,h]anthracene as the reference compound (Fent, 2003).

Advantages to the TEF procedure (Table 15) include the relative ease of determining EROD activity in hepatic cell cultures (Fent, 2003). Similar *in vitro* assays may be developed for other receptor mediated responses in the future, allowing health risk estimation for understudied congeners. TEFs predict health effects well for the species in which they are developed (George *et al.*, 2003). Contamination of congeners can be handled as if an equivalent concentration of the reference compound is present at the site, simplifying risk characterization but not uncertainty assessment (Safe, 1994; Schoeny and Margoshes, 1989; Teuschler and Hertzberg, 1995).

The major disadvantage to using the TEF procedure in ERA is the species specificity of the response on which the equivalency is based. TEFs are assigned to a class of receptors (e.g., mammals, piscivore fish) but are based on the response of a single species to the chemical (George *et al.*, 2003; Putzrath, 1997). It is considered acceptable for different mammals to differ by an order of magnitude in response to a contaminant; responses of non-mammals differ even more greatly. Yet, when TEF data are unavailable for different receptor classes, mammalian TEFs may be used (Sanderson and van den Berg, 1999).

Another disadvantage of the TEF procedure is common to all component approaches. Components of a whole mixture that are present below detection limits can have unexpected effects on toxicity. Induction equivalents from whole mixture tests with fish hepatocytes have resulted in CYP1a levels 4 to 112 times the activity predicted from dose addition of the measured components concentrations and IEFs (Fent, 2003). TEFs can underpredict toxicity for a different reason, as well. The TEF procedure only accounts for response mitigated by a specific receptor. Other modes of action are not assessed and may contribute to the overall toxicity of the whole mixture (Sanderson and van den Berg, 1999).

Table 15. Advantages and Disadvantages of the TEF Procedure

Advantages	Disadvantages
• Predicts well for chemicals with same mode of action • <i>In vitro</i> data easily gathered • Toxic equivalents stated in concentration units of reference compound • Allows risk quantitation for congeners without sufficient toxicity databases	• Interactions not incorporated • Data insufficient to verify dose-response curves and mode of action • Needs detailed knowledge of mixture components and concentrations • Species specific TEFs applied to whole classes of receptors • Undetected chemicals can add to toxicity • Ignores additional sensitive endpoints and modes of action

Risk Procedures for Whole Mixture Approaches

Whole mixture approaches are generally reserved for Tier 2 or 3 investigations, dictated by the need for site specific toxicity data (Wentsel *et al.*, 1994; 1996). Whole mixture approaches employ biological assays to measure the combined effects of complex environmental mixture exposures (Ankley and Mount, 1996). A component approach is often impractical due to the number of poorly characterized components present at hazardous waste sites or in effluents released to the environment. Only 100 to 150 priority pollutants are routinely screened under the Clean Water Act (Smolders *et al.*, 2003). At RCRA and Superfund sites, contaminant screens are generally limited to U.S. EPA priority pollutants (Barron and Holder, 2003). Toxicity and interactions can be caused by non-routine contaminants or priority chemicals can be below detection limits (Gardner *et al.*, 1998). The complexity of mixtures at sites makes whole mixtures approaches the preferred method of assessing toxicity in baseline and Tier 3 ERAs (U.S. EPA, 1997a; 1998).

Whole mixture testing encompasses two main types of evaluations, laboratory assays and field studies. There is no standard protocol for investigating site mixtures but it is generally accepted that an integrated approach should be used, combining more than one assay and potentially more than one type of test to develop lines of evidence for the ERA (Twerdok *et al.*,

1997). The type or types of assays chosen depend on the toxic effect being studied or the goal of the test (e.g., relevant test species reproduction assay to infer long term viability of site species population) (Ankley and Mount, 1996; Van Leeuwen *et al.*, 1996).

Laboratory evaluations include *in vitro* (cellular endpoints) and *in vivo* (whole animal endpoints) bioassays. The assays are performed using water, sediment or soil taken from the contaminated site to project environmental impacts of the whole mixture. Mixture contents vary temporally and spatially; multiple sampling points and times from a site are necessary. It is assumed that the testing of surrogate species in site water or on site substrates is indicative of likely effects in receptors of concern (Chapman, 2000; U.S. EPA, 1997a; Wentzel *et al.*, 1994; 1996).

An integrated assessment using *in vivo* and/or *in vitro* assays should include multiple tests measuring acute and chronic endpoints, as well as a chemical analysis (Twerdok *et al.*, 1997). The whole effluent toxicity (WET) testing battery is an example of an integrated *in vivo* approach to whole mixture toxicity. These standardized protocols are designed for measuring the toxicity of aqueous effluents to freshwater, marine and estuarine organisms prior to release in the environment. The approved acute tests include water flea (*Ceriodaphnia dubia*), inland silverside fish (*Menidia beryllina*), fathead minnow (*Pimephales promelas*) and sheepshead minnow (*Cyprinodon variegates*) survival bioassays. Chronic assays include algal (*Selenastrum capricornutum*) growth, *Ceriodaphnia dubia* reproduction, mysid shrimp (*Mysidopsis bahia*) growth and fecundity, and fish (inland silverside, fathead minnow or sheepshead minnow) larvae survival and growth tests. At least three test species from different phyla must be used in order to receive an effluent discharge permit (Norberg-King *et al.*, 1992; U.S. EPA, 2002a).

Laboratory test conditions are controlled; temperature, climate, water conditions and diet do not fluctuate as they do in the environment. Control reduces variability; however, inter- and intra-laboratory differences in results commonly reach a factor of two. Controlled test conditions may increase conservativeness of the results in some ways. Mobile organisms are not

permitted to avoid exposure to the mixture as they might in the environment. Natural degradation processes such as photolysis do not occur. Test organisms are not given the opportunity to adapt or acclimate. *In vivo* tests may underpredict toxicity in other situations where a site physical condition (e.g., food scarcity, low dissolved oxygen level) augments the toxicity of the chemical mixture. Test conditions are not environmentally realistic; effects seen in the field may not concur with laboratory results for the site mixture (Chapman, 2000; Fent, 2003).

The second type of whole mixture toxicity approaches, field studies, includes *in situ* bioassays and biomonitoring studies. *In situ* studies utilize cages of organisms exposed at different sampling points at the site along spatial gradients of contamination. These organisms experience ambient conditions, including temperature extremes, food restrictions and disease, and are assumed to react to contamination in a similar fashion as site species (Foran and Ferenc, 1997; Smolders *et al.*, 2003; U.S. EPA, 1997a). Organisms are generally laboratory-raised species but may be locally prevalent organisms gathered from a reference site; wild-type organisms acclimate to natural surroundings without some of the shock seen in lab raised individuals. *In situ* studies allow comparison of two sites where similar species may not be naturally found (Smolders *et al.*, 2003).

Biomonitoring studies measure some endpoint or effect in the existing site ecosystem (Van Leeuwen *et al.*, 1996). Measurements in indigenous animals increase ecological relevance (Carlson *et al.*, 2003). Endpoints include algae or plant density, benthic macroinvertebrate surveys, fish indices, plant density and soil invertebrate diversity indices. Measurements may be more function, rather than population, related (e.g., algae carbon uptake or photosynthesis rate) (Fent, 2003). Biomonitoring may also include the gathering of exposure biomarkers; site mollusks, earthworms, birds or mammal populations can be sampled for tissue residues to determine if exposure pathways are complete or if bioaccumulation is occurring

(Carlson *et al.*, 2003; Menzie *et al.*, 1992; Smolders *et al.*, 2003). The measurement endpoint is assumed to infer a health effect to the population of concern.

Both types of field studies provide evidence of impacts currently occurring at the site but are not useful for predicting toxicity if the exposure pathway has not yet been completed (i.e., chemicals have not yet reached receptor species) (U.S. EPA, 1997a; Van Leeuwen *et al.*, 1996). Both accommodate temporal changes in contaminants, such as storm events or effluent surges (Smolders *et al.*, 2003) and spatial distribution of effects (Menzie *et al.*, 1992). Field studies provide measures of bioaccumulation; tissue residues gathered from *in situ* or indigenous organisms indicate body burden (Chapman, 2000). The rate of accumulation is easily calculated using *in situ* organisms, as their exposure period is known (Menzie *et al.*, 1992; Smolders *et al.*, 2003). Field studies, however, do not provide good information on the causes of effects seen (Van Leeuwen *et al.*, 1996).

Using whole mixture data in risk assessments is not as straight forward as is the process for component approaches. Laboratory tests are used in ERA to determine potential toxicity of a site mixture. Field studies support or undermine these conclusions. However, whole mixture tests do not provide information on the causes of toxicity at a site unless further studies are undertaken (Ankley and Mount, 1996). Causality must be established in order for the site manager to make decisions regarding remediation. Causality is the relationship between a stressor or multiple stressors and adverse response(s). Uncertainty in the ERA will be very high if a strong causal relationship is not established. Toxicity identification evaluations (TIEs) isolate components of the mixture that have toxic effects and their relative contributions to the whole mixture toxicity. Fractions of the site mixture are either separated from the whole or reconstructed from known, suspected components. Laboratory toxicity tests are then run on the submixture. During a TIE, mixture component toxicity results can be compared to the whole mixture effects, depending on data availability. TIEs may be used to estimate mixture component concentrations that should result in an acceptable level of impact (Foran and

Ferenc, 1997; U.S. EPA, 1998). Whole mixture studies and the ensuing TIE analyses are often expensive and technically challenging (Ankley and Mount, 1996; Fairbrother, 2003). However, their use is considered worthwhile for large-scale assessments where significant cleanup costs are at stake (Fairbrother, 2003).

Whole mixture approaches avoid three major disadvantages of component approaches, missed or below detection chemicals, chemical to chemical interaction and bioavailability (Table 16) (Ankley and Mount, 1996; Fent, 2003; Smolders *et al.*, 2003; Van Leeuwen *et al.*, 1996). Site water and substrate samples are assumed to contain all site contaminants in their relative concentrations and bioavailability states. Bioavailability of chemicals and metals is dependent on chemical and physical processes in the medium. Only the bioavailable portion of a toxicant can result in adverse effects (Fent, 2003). For metals, oxidation state determines availability to receptors. For example, absorption of orally administered trivalent chromium is about 1% while approximately 10% of hexavalent chromium is absorbed (Schoof, 2003; U.S. EPA, 1997a). Both types of whole mixture testing take some aspects of bioavailability into account; however, disturbances (mixing) of sediments and soils during grab or composite sampling can alter toxicity in laboratory evaluations (Menzie *et al.*, 1992; Smolders *et al.*, 2003). Careful evaluation of the site may indicate invertebrate populations on the surface of the sediment or soil only, not deeper where toxins are undiluted (Menzie *et al.*, 1992). Test species used in laboratory or *in situ* assays may be sensitive to background metals concentrations to which indigenous organisms have adapted; these assays could not be used for evaluation at the site (Chapman, 2000).

Unknown toxicants can have substantial impact on site toxicity. Significant toxicity can occur in effluents at dilutions an order of magnitude lower than predicted by component methods, presumably due to undetected chemicals in the mixture (Smolders *et al.*, 2003). Groundwater used for toxicity testing in Medaka fish (*Oryzias latipes*) resulted in liver carcinogenicity. Trichloroethylene was the only reportable contaminant but unidentified and

undifferentiated peaks were seen on the gas chromatogram; these chemicals or their interactions with TCE are assumed to be responsible as TCE alone is not a carcinogen in Medaka (Gardner *et al.*, 1998). Although whole mixture approaches provide a way to assess unknown toxicants, incomplete knowledge of toxic components can complicate establishing causality and determining cleanup levels (Smolders *et al.*, 2003; Van Leeuwen *et al.*, 1996).

As for component approaches, extrapolation between test species and site receptors in whole mixture assays adds uncertainty to the approach (Fent, 2003). Extrapolations of chemical toxicity levels are frequently based solely on bodyweight without consideration of additional physiological differences (Fairbrother, 2003). In laboratory assays or *in situ* tests, test species are inbred for reduced variation (Carlson *et al.*, 2003; U.S. EPA, 1997a); the most sensitive site species can not be raised in a laboratory. Assay species may only be distantly related to the site species; toxicity tests vary significantly between members of the same genus (Chapman, 2000). Species extrapolation can also be a concern when using biomonitoring indices. Indexed groups may not be representative of the species of concern; invasive biomarkers can not be collected from threatened or endangered species. Assumptions on the health status of entire ecosystems rest on incomplete toxicology in a few species (Smolders *et al.*, 2003; Van Leeuwen *et al.*, 1996).

Both laboratory and field studies require local reference sites used to determine baseline survival rates, background contamination level or body burden, and biotic index scores. These sites can be difficult to find. Mixture components must be well characterized in order to determine if the reference site is relatively unpolluted. If such a site is not available, the least stressed site may be used or reference values may be derived from data distributions based on multiple distant sites (Foran and Ferenc, 1997; Sanderson and van den Berg, 1999).

Table 16. Advantages and Disadvantages of Whole Mixture Laboratory and Field Toxicity Testing Procedures

Advantages	Disadvantages
• Incorporates components, interactions, mechanisms of action • Accommodates real life exposures • Accounts for bioavailability • LAB: Controlled conditions help reduce variability • FIELD: Provides bioaccumulation data • FIELD: Incorporates temporal and spatial differences	• Species to species extrapolation increases uncertainty • Reference sites difficult to find • Lack data on causal relationships • TIEs expensive, challenging and time consuming • LAB: Controlled conditions prevent environmental mixture composition changes • LAB: Not useful for background metal exposures

Apparent Effects Threshold Procedure

The apparent effects threshold (AET) procedure represents a specific use of whole mixture toxicity tests for setting sediment quality criteria. The main assumption behind the AET procedure is that, above a certain threshold, concentrations of common site contaminants in sediments will result in significant biological effects. A dose-response relationship is assumed not only for benthic organisms, but also fish. Sediments serve as a sink and a source to overlying water for persistent contaminants; sediment contaminant load is assumed to be the predominant source of toxicity (Alden and Rule, 1992; Barrick *et al.*, 1989; MacDonald *et al.*, 2000; Swartz, 1999; Wentsel *et al.*, 1994; 1996). The AET approach was developed for Puget Sound and is designed for large ecosystems where similar contaminants are likely to occur at multiple sampling points (MacDonald and Ingersoll, 2002).

AETs are established by first collecting sediments from multiple sampling points for chemical analysis, expressed in dry weight normalized terms. Laboratory assays using the contaminated sediment and benthic or water column organisms are performed. Microtox bacteria acute toxicity, amphipod mortality and oyster larvae abnormalities are examples of

assays used in setting AETs. Benthic biomonitoring indices can also be used to determine sampling point toxicity. AETs are set separately for different taxa as organisms have differential sensitivity to contaminants. To determine AETs, statistically significant ($p \leq 0.05$) effects must be observed at some sampling points, as compared to a reference site, and chemical components must be above analytical detection. For every compound in the chemical analysis, effects are categorized as significant or non-significant for each sampling point. The highest concentration at a sampling point which did not result in significant toxicity becomes the threshold concentration for that chemical and species. When AETs are assigned for prevalent chemicals and multiple species, the AETs are verified with data from additional sampling points before they are used throughout the ecosystem. The toxicity of a site, as predicted by exceeding one or more AETs, should correlate with biomonitoring indices or laboratory tests made with sediment from the sampling point. The AET procedure has been found to predict the toxicity of 90 to 94% of new sampling sites (Barrick *et al.*, 1989; MacDonald and Ingersoll, 2002; MacDonald *et al.*, 2000). AETs can be considered site specific values, albeit the sites for which they are developed are really entire ecosystems. Site specific validation of AETs is recommended if they are used outside the originating site (Barrick *et al.*, 1989).

The advantage to an AET is that it is based on an empirical relationship between contaminant concentrations and laboratory or field toxicity data (Table 17). This relationship is verified with independent data prior to use (Alden and Rule, 1992). The main disadvantage is that measured chemical concentrations are not the sole reasons for biological responses seen at sampling sites. Unmeasured components can change toxicity; physical stressors and biological interactions between species can alter biomonitoring index results (Alden and Rule, 1992; Barrick *et al.*, 1989; Wentsel *et al.*, 1994; 1996). Determining causality for observed toxicity is difficult (MacDonald and Ingersoll, 2002). The AET procedure is inherently prone to false negatives. A chemical AET is likely to increase with each new sampling point where no significant toxicity is found. For this reason, AETs have been found to correctly identify highly

toxic sites but to overlook sites where lower levels of contamination occur, even though these lesser contaminated sites are shown to have significant toxicity when bioassays are performed (MacDonald and Ingersoll, 2002).

Table 17. Advantages and Disadvantages of the AET Procedure

Advantages	Disadvantages
• Based on empirical relationship between site contaminants and toxicity	• Measured contaminants not solely responsible for site toxicity • Determining causality difficult • Prone to false negatives • Reference sites difficult to find

CONCLUSIONS

This paper reviews the toxicological principles used in mixtures risk assessment. Human and ecological risk assessments have been handled separately due to the difference in objectives between the types of assessments. ERA focuses on population effects while HHRA attempts to prevent adverse effects in sensitive individuals (Suter, 2004; Wireman, personal communication, 2004). Because the objectives are dissimilar, the direction of mixtures risk assessment also differs between disciplines. Human health risk assessments are currently putting more emphasis on interactions methods based on component-type assessments. This is seen in the newest statistical method, the interaction HI (U.S. EPA, 2000), and the continued interest in mechanism-based methods including PBPK/PD modeling (Robinson and MacDonell, 2004). While whole mixture toxicity is considered the first and best method (U.S. EPA, 2000), the minimal database of toxicity tests for HHRA is too expensive to undertake on a site-by-site or mixture-by-mixture basis. This contrasts with ecological risk assessment, where the emphasis remains on whole mixtures effects. Ecological sites are often larger, sometimes incorporating entire ecosystems (U.S. EPA, 1997b). An integrated battery of studies, frequently

utilizing invertebrates, is relatively cheaper, faster and more attainable on a site-by site basis. Component approaches, currently limited to the standard HI and the related TEFs, are even more complicated for ecological receptors than human. Hazard quotients are summed not only for each mode of action, as in HHRA (U.S. EPA, 2000), but also for each type of receptor or related receptor class (Wentsel *et al.*, 1994; 1996). Relatively scarce toxicity data and even fewer interactions data exist for ecological receptors compared to the number of studies designed to assess human health risk.

Yet, HHRAs and ERAs of mixtures have common aspects. Both have component and whole mixtures approaches to dealing with mixtures risk. All these approaches have their advantages and disadvantages. No one approach will be sufficient for all types of sites and assessment needs (Seed *et al.*, 1995). Both disciplines sorely need more research into mixtures toxicity; lack of toxicity and interactions data is the most frequently listed disadvantage for all the approaches. Studies must have appropriate statistical design in order for the data to be used effectively (Simmons, 1995). Mixture HHRAs and ERAs may be misunderstood as having detailed knowledge of interactions toxicity and, therefore, the potential health risk at the site. Following each assessment must be a detailed analysis of uncertainty, for which there are many contributing factors (U.S. EPA, 2000).

For all the shortcomings of mixtures risk assessments, there are advantages to performing them outside of regulatory requirement. A mixtures assessment, regardless of approach, will do a better job of reflecting site risk as opposed to a single chemical or multiple separate chemical assessments. This assessment provides the site manager with more information on which to base decisions and prioritize problems. Information can mean reduced uncertainty and increased confidence in the risk estimate (ILSI, 1999).

Future directions of human and ecological mixtures assessments are also similar. Biomarkers are already used in ecological assessments to determine if exposure pathways for chemicals are complete and if bioaccumulation is occurring (Carlson *et al.*, 2003; Menzie *et al.*,

1992; Smolders *et al.*, 2003). Occupational assessments of mixtures are progressing more into the use of biomarkers (Viau, 2002). Occupational mixtures are rarely measured fully, so component approaches to mixtures risk is limited to the known chemicals in the workplace exposure. Biomarkers show promise in intervention strategies to prevent overt toxicity by measuring early cellular responses. Of course, the value of a particular biomarker is limited by the knowledge of its relation to toxicity and the specificity of its response (NORA, 2004). Biomarkers reflect the exposures from the total environment, not just the work environment. For this reason, biomarkers of metabolic enzymes such as CYP1a are not as useful in humans as in ecological receptors; CYP1a is induced by common personal "exposures" such as ethanol (Raucy, 1995).

Genomic studies will play a part in both ecological and human mixtures assessments in the future. Genomics is the study of changes in cellular expression in response to their environment (Travis *et al.*, 2003). Gene arrays allow simultaneous evaluation of expression levels for up to 10,000 genes (Teuschler *et al.*, 2001). These arrays may be used for screening cellular responses to single chemicals or mixtures and determining the mode(s) of action involved in their toxicity; mixtures responses may be compared to individual component responses to determine if interaction exists and the magnitude of the difference in response can be easily calculated (Kleinjans, 2003; Robinson and MacDonell, 2004; Travis *et al.*, 2003). Genomics may be able to detect effects at low, environmentally relevant concentrations and determine if interactions truly are not present at those levels, as is often assumed (ATSDR, 2001a; Carpy *et al.*, 2000; Feron *et al.*, 2002). Hazard screening of new chemicals or mixtures could be prioritized through genomic assays; using comparisons of gene expression from known toxicants, toxicity of new components would be estimated (Kleinjans, 2003; Teuschler *et al.*, 2002; Thomas *et al.*, 2002). Genomics may also provide a simple method of estimating TEFs for congeners once the cellular response to the reference chemical is well characterized. These assays have the potential for providing cellular dynamic information useful for human

PBPD modeling. In turn, PBPK/PD models developed in parallel with genomics studies can help link *in vitro* cellular responses to environmental exposure levels (Andersen and Dennison, 2004). Gene arrays could have the capability of determining the loss in microbial diversity in contaminated soil compared to uncontaminated sites (Travis *et al.*, 2003).

This review examines the toxicological approaches currently used for mixtures risk assessments in human and ecological receptors. Today, quantitative and semi-quantitative methods exist for predicting risks of mixtures and sometimes interactions of components. Due to data restrictions, qualitative approaches are used when necessary (U.S. EPA, 2000; 2002a). Few exposures are to single chemicals (U.S. EPA, 1986; Viau, 2002). More research and innovative approaches are necessary to continue to improve methods for predicting quantitative risks to the ubiquitous mixtures.

REFERENCES

- Abebe, W. 2002. Herbal medication: Potential for adverse interactions with analgesic drugs. *J. Clin. Pharm. Therapeut.* 27:391-401.
- Alden, R. W. and Rule, J. H. 1992. Uncertainty and sediment quality assessments. 2. Effects of correlations between contaminants on the interpretation of apparent effects threshold data. *Environ. Toxicol. Chem.* 11:645-651.
- Altenburger, R., Nendza, M., and Schuurmann, G. 2003. Mixture toxicity and its modeling by quantitative structure-activity relationships. *Environ. Toxicol. Chem.* 22:1900-1915.
- Andersen, M. E. and Dennison, J. E. 2004. Mechanistic approaches for mixture risk assessments—present capabilities with simple mixtures and future directions. *Environ. Toxicol. Pharmacol.* 16:1-11.
- Andersen, M. E., Gargas, M. L., Clewell, H. J., III, and Severyn, K. M. 1987. Quantitative evaluation of the metabolic interactions between trichloroethylene and 1,1-dichloroethylene in vivo using gas uptake methods. *Toxicol. Appl. Pharmacol.* 89:149-157.
- Ankley, G. T. and Mount, D. R. 1996. Retrospective analysis of the ecological risk of contaminant mixtures in aquatic sediments. *Human. Ecol. Risk Assess.* 2:434-440.
- ATSDR. 1998. Toxicological profile for jet fuels JP-5 & JP-8. Agency for Toxic Substances and Disease Registry, U.S. Department of Health and Human Services, Public Health Service, Atlanta, GA. DD-97/04. www.atsdr.cdc.gov/toxpro2.
- ATSDR. 2001a. Draft Guidance Manual for the Assessment of Joint Toxic Action of Chemical Mixtures. Agency for Toxic Substances and Disease Registry, U.S. Department of Health and Human Services, Public Health Service, Atlanta, GA.
- ATSDR. 2001b. Draft interaction profile for cesium, cobalt, polychlorinated biphenyls, strontium and trichloroethylene. Agency for Toxic Substances and Disease Registry, U.S. Department of Health and Human Services, Public Health Service, Atlanta, GA.
- ATSDR. 2002. Draft interaction profile for cyanide, fluoride, nitrate and uranium. Agency for Toxic Substances and Disease Registry, U.S. Department of Health and Human Services, Public Health Service, Atlanta, GA.
- Barrick, R., Beller, H., Becker, S., and Ginn, T. 1989. Use of the apparent effects threshold approach (AET) in classifying contaminated sediments. In: *Contaminated Marine Sediments: Assessment and Remediation*. C. o. C. M. S. National Research Council, ed. Washington, D.C.: National Academy Press. p. 64-77.
- Barron, M. G. and Holder, E. 2003. Are exposure and ecological risks of PAHs underestimated at petroleum contaminated sites. *Human. Ecol. Risk Assess.* 9:1533-1545.
- Borgert, C. J., LaKind, J. S., and Witorsch, R. J. 2003. A critical review of methods for comparing estrogenic activity of endogenous and exogenous chemicals in human milk and infant formula. *Environ. Health Perspect.* 111:1020-1036.
- Borgert, C. J., Price, B., Wells, C. S., and Simon, G. S. 2001. Evaluating chemical interaction studies for mixture risk assessment. *Human. Ecol. Risk Assess.* 7:259-306.
- Brown, D. G., Knightes, C. D., and Peters, C. A. 1999. Risk assessment for polycyclic aromatic hydrocarbon NAPLs using component fractions. *Environ. Sci. Technol.* 33:4357-4363.
- Burton, G. A., Chapman, P. M., and Smith, E. P. 2002. Weight-of-evidence approaches for assessing ecosystem impairment. *Human. Ecol. Risk Assess.* 8:1657-1673.
- Cahill, T. M., Cousins, I., and Mackay, D. 2003. General fugacity-based model to predict the environmental fate of multiple chemical species. *Environ. Toxicol. Chem.* 22:483-493.
- Calabrese, E. J. 1995. Toxicological consequences of multiple chemical interactions: A primer. *Toxicology.* 105:121-135.

- Carlson, R. I., Willson, J. A., Lochmiller, R. L., Janz, D. M., Schroder, J. L., and Basta, N. T. 2003. Ecotoxicological risks associated with land treatment of petrochemical wastes. II. Effects on hepatic phase I and phase II detoxification enzymes in cotton rats. *J. Toxicol. Environ. Health A*. 66:327-343.
- Carpenter, D. O., Arcaro, K. F., Bush, B., Niemi, W. D., Pang, S., and Vakharia, D. D. 1998. Human health and chemical mixtures: An overview. *Environ. Health Perspect.* 106:1263-1270.
- Carpenter, D. O., Arcaro, K., and Spink, D. C. 2002. Understanding the human health effects of chemical mixtures. *Environ. Health Perspect.* 110:25-42.
- Carpy, S. A., Kobel, W., and Doe, J. 2000. Health risk of low-dose pesticides mixtures: A review of the 1985-1998 literature on combination toxicology and health risk assessment. *J. Toxicol. Environ. Health B Crit. Rev.* 3:1-25.
- Cassee, F. R., Groten, J. P., van Bladeren, P. J., and Feron, V. J. 1998. Toxicological evaluation and risk assessment of chemical mixtures. *Crit. Rev. Toxicol.* 28:73-101.
- Chapman, P. M. 2000. Whole effluent toxicity testing - Usefulness, level of protection, and risk assessment. *Environ. Toxicol. Chem.* 19:3-13.
- Clark, L. H., Setzer, R. W., and Barton, H. A. 2004. Framework for evaluation of physiologically-based pharmacokinetic models for use in safety or risk assessment. *Risk Anal.* 24:1697-1717.
- Clay, J. and Harris, M. E. 2002. Risk-based corrective action of hydrocarbon contamination at a former major urban petroleum storage site in the UK. *Soil Sed. Contam.* 11:701-718.
- Conolly, R. B. 2001. Biologically motivated quantitative models and the mixture toxicity problem. *Toxicol. Sci.* 63:1-2.
- Craig, D. K., Baskett, R. L., Davis, J. S., Dukes, L., Hansen, D. J., Petrocchi, A. J., Powell, T. J., Sutherland, P. J., and Tuccinardi, T. E. 1999. Recommended default methodology for analysis of airborne exposures to mixtures of chemicals in emergencies. *Appl. Occup. Environ. Hyg.* 14:609-617.
- Deneer, J. W. 2000. Toxicity of mixtures of pesticides in aquatic systems. *Pest Management Science.* 56:516-520.
- Dennison, J. E., Andersen, M. E., and Yang, R. S. 2003. Characterization of the pharmacokinetics of gasoline using PBPK modeling with a complex mixtures chemical lumping approach. *Inhal. Toxicol.* 15:961-986.
- Dennison, J. E., Andersen, M. E., Dobrev, I. D., Mumtaz, M. M., and Yang, R. S. H. 2004. PBPK modeling of complex hydrocarbon mixtures: Gasoline. *Environ. Toxicol. Pharmacol.* 16:107-119.
- Dobrev, I. D., Andersen, M. E., and Yang, R. S. H. 2002. *In silico* toxicology: Simulating interaction thresholds for human exposure to mixtures of trichlorethylene, tetrachlorethylene, and 1,1,1-trichloroethane. *Environ. Health Perspect.* 110:1031-1039.
- Donnelly, K. C., Safe, S. H., Randerath, K., and Randerath, E. 1995. Bioassay-based risk assessment of complex-mixtures. *J. Hazard. Materials.* 41:341-350.
- Dressler, V., Muller, G., and Suhnel, J. 1999. CombiTool - A new computer program for analyzing combination experiments with biologically active agents. *Comput. Biomed. Res.* 32:145-160.
- Durkin, P., Hertzberg, R., Stiteler, W., and Mumtaz, M. 1995. The identification and testing of interaction patterns. *Toxicol. Lett.* 79:251-264.
- Dyer, R. S. 2002. Chemical mixtures and health effects at Superfund sites. *Int. J. Hyg. Environ. Health.* 205:149-153.
- Dyer, S. D., White-Hull, C. E., and Shephard, B. K. 2000. Assessments of chemical mixtures via toxicity reference values overpredict hazard to Ohio fish communities. *Environ. Science Technol.* 34:2518-2524.
- ECETOC. 2001. Aquatic toxicity of mixtures (abstract). European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium. Technical Report No. 80.

- El-Masri, H. A., Mumtaz, M. M., and Yushak, M. L. 2004. Application of physiologically-based pharmacokinetic modeling to investigate the toxicological interaction between chlorpyrifos and parathion in the rat. *Environ. Toxicol. Pharmacol.* 16:57-71.
- El-Masri, H. A., Reardon, K. F., and Yang, R. S. H. 1997. Integrated approaches for the analysis of toxicologic interactions of chemical mixtures. *Crit. Rev. Toxicol.* 27:175-197.
- Escher, B. I. and Hermens, J. L. M. 2002. Modes of action in ecotoxicology: their role in body burdens, species sensitivity, QSARs, and mixture effects. *Environ. Science Technol.* 36:4201-4217.
- Fairbrother, A. 2003. Lines of evidence in wildlife risk assessments. *Human. Ecol. Risk Assess.* 9:1475-1491.
- Faust, M., Altenburger, R., Backhaus, T., Bodeker, W., Scholze, M., and Grimme, L. H. 2000. Predictive assessment of the aquatic toxicity of multiple chemical mixtures. *J. Environ. Qual.* 29:1063-1068.
- Fent, K. 2003. Ecotoxicological problems associated with contaminated sites. *Toxicol. Lett.* 140-141:353-365.
- Feron, V. J. and Groten, J. P. 2002. Toxicological evaluation of chemical mixtures. *Food Chem. Toxicol.* 40:825-839.
- Feron, V. J., Cassee, F. R., Groten, J. P., vanVliet, P. W., and van Zorge, J. A. 2002. International issues on human health effects of exposure to chemical mixtures. *Environ. Health Perspect.* 110:893-899.
- Fisher, J., Lumpkin, M., Boyd, J., Mahle, D., Bruckner, J. V., and El-Masri, H. A. 2004. PBPK modeling of the metabolic interactions of carbon tetrachloride and tetrachloroethylene in B6C3F1 mice. *Environ. Toxicol. Pharmacol.* 16:93-105.
- Foran, J. A. and Ferenc, S. A. 1997. Multiple Stressors in Ecological Risk and Impact Assessment: Proceedings from the Pellston Workshop on Multiple Stressors in Ecological Risk and Impact Assessment. Pensacola, FL: SETAC Press. pp. WSU QH541.15.R57 P4485 1997.
- Gardner, H. S., Brennan, L. M., Toussaint, M. W., Rosencrance, A. B., Boncavage-Hennessey, E. M., and Wolfe, M. J. 1998. Environmental complex mixture toxicity assessment. *Environ. Health Perspect.* 106:1299-1305.
- George, T. K., Liber, K., Solomon, K. R., and Sibley, P. K. 2003. Assessment of the probabilistic ecological risk assessment-toxic equivalent combination approach for evaluating pesticide mixture toxicity to zooplankton in outdoor microcosms. *Arch. Environ. Contam. Toxicol.* 45:453-461.
- Grimme, L. H., Faust, M., Boedeker, W., and Altenburger, R. 1996. Aquatic toxicity of chemical substances in combination: Still a matter of controversy. *Human. Ecol. Risk Assess.* 2:426-433.
- Groten, J. P. 2000. Mixtures and interactions. *Food Chem. Toxicol.* 38:S65-S71.
- Haas, C. N., Cidambi, K., Kersten, S., and Wright, K. 1996. Quantitative description of mixture toxicity: Effect of level of response on interactions. *Environ. Toxicol. Chem.* 15:1429-1437.
- Haddad, S. and Krishnan, K. 1998. Physiological modeling of toxicokinetic interactions: Implications for mixture risk assessment. *Environ. Health Perspect.* 106:1377-1384.
- Haddad, S., Beliveau, M., Tardif, R., and Krishnan, K. 2001. A PBPK modeling-based approach to account for interactions in the health risk assessment of chemical mixtures. *Toxicol. Sci.* 63:125-131.
- Health Council of the Netherlands. 2002. Exposure to combinations of substances: A system for assessing health risks. Health Council of the Netherlands, The Hague. Publication no. 2002/05.
- Henschler, D., Bolt, H. M., Jonker, D., Pieters, M. N., and Groten, J. P. 1996. Experimental designs and risk assessment in combination toxicology: panel discussion. *Food Chem. Toxicol.* 34:1183-1185.

- Hertzberg, R. C. and MacDonell, M. M. 2002. Synergy and other ineffective mixture risk definitions. *Sci. Total Environ.* 288:31-42.
- Hutcheson, M. S., Pedersen, D., Anastas, N. D., Fitzgerald, J., and Silverman, D. 1996. Beyond TPH: Health-based evaluation of petroleum hydrocarbon exposures. *Regul. Toxicol. Pharmacol.* 24:85-101.
- ILSI. 1999. A framework for cumulative risk assessment. An ILSI Risk Science Institute Workshop Report. ILSI Press, International Life Science Institute, Washington, D.C.
- Johnson, B. L. and DeRosa, C. T. 1995. Chemical mixtures released from hazardous waste sites: Implications for health risk assessment. *Toxicology.* 105:145-156.
- Johnson, Mark S. 2004. Personal communication (e-mail) with Mark S. Johnson, Ph.D., DABT, Health Effects Research Program, Directorate of Toxicology, U.S. Army Center for Health Promotion and Preventative Medicine, Aberdeen Proving Ground, MD. 26 Jan 2004.
- Kelsey, J. W. and Alexander, M. 1997. Declining bioavailability and inappropriate estimation of risk of persistent compounds. *Environ. Toxicol. Chem.* 16:582-585.
- Kleinjans, J. C. S. 2003. Principles in toxicological risk analysis. *Toxicol. Lett.* 140-141:311-315.
- Kodell, R. L., Ahn, J., Chen, J. J., Springer, J. A., Barton, C. N., and Hertzberg, R. C. 1995. Upper bound risk estimates for mixtures of carcinogens. *Toxicology.* 105:199-208.
- Krishnan, K., Clewell, H. J., III, and Andersen, M. E. 1994. Physiologically based pharmacokinetic analyses of simple mixtures. *Environ. Health Perspect.* 102 Suppl 9:151-155.
- Krishnan, K., Haddad, S., Beliveau, M., and Tardif, R. 2002. Physiological modeling and extrapolation of pharmacokinetic interactions from binary to more complex chemical mixtures. *Environ. Health Perspect.* 110:989-994.
- Krishnan, K., Paterson, J., and Williams, D. T. 1997. Health risk assessment of drinking water contaminants in Canada: The applicability of mixture risk assessment methods. *Regul. Toxicol. Pharmacol.* 26:179-187.
- Lang, L. 1995. Strange brew: Assessing risk of chemical mixtures. *Environ. Health Perspect.* 103:142-145.
- Lebron, C. 2003. Enhanced natural attenuation of commingled plumes. *Remedial Project Manager News.* Winter:6-9.
- Liao, K. H., Dobrev, I. D., Dennison, J. E., Andersen, M. E., Reisfeld, B., Reardon, K. F., Campaign, J. A., Wei, W., Klein, M. T., Quann, R. J., and Yang, R. S. H. 2002. Application of biologically based computer modeling to simple or complex mixtures. *Environ. Health Perspect.* 110:957-963.
- Logan, D. T. and Wilson, H. T. 1995. An ecological risk assessment method for species exposed to contaminant mixtures. *Environ. Toxicol. Chem.* 14:351-359.
- MA DEP. 2002. Characterizing risks posed by petroleum contaminated sites: Implementation of the MA DEP VPH/EPH approach. Final policy. Massachusetts Department of Environmental Protection, Commonwealth of Massachusetts, Boston, MA. Policy #WSC-02-411.
- MA DEP. 2003. Final. Updated petroleum hydrocarbon fraction toxicity values for the VPH/EPH/APH methodology. Massachusetts Department of Environmental Protection, Commonwealth of Massachusetts, Boston, MA. Policy #WSC-02-411.
- MacDonald, D. D. and Ingersoll, C. G. 2002. A guidance manual to support the assessment of contaminated sediments in freshwater ecosystems. Volume III: Interpretation of the results of sediment quality investigations. U.S. Environmental Protection Agency, Chicago, IL. EPA-905-B02-001-C.
- MacDonald, D. D., Dipinto, L. M., Field, J., Ingersoll, C. G., Long, E. R., and Swartz, R. C. 2000. Development and evaluation of consensus-based sediment effect concentrations for polychlorinated biphenyls. *Environ. Toxicol. Chem.* 19:1403-1413.
- Malich, G., Markovic, B., and Winder, C. 1998. Human cell line toxicity of binary and ternary chemical mixtures in comparison to individual toxic effects of their components. *Arch. Environ. Contam. Toxicol.* 35:370-376.

- Mauderly, J. L. 1993. Toxicological approaches to complex mixtures. *Environ. Health Perspect.* 101:155-165.
- Menzie, C. A., Burmaster, D. E., Freshman, J. S., and Callahan, C. A. 1992. Assessment of methods for estimating ecological risk in the terrestrial component - A case-study at the Baird and McGuire Superfund Site in Holbrook, Massachusetts. *Environ. Toxicol. Chem.* 11:245-260.
- Mumtaz, M. M. and Durkin, P. R. 1992. A weight-of-evidence approach for assessing interactions in chemical mixtures. *Toxicol. Ind. Health.* 8:377-406.
- Mumtaz, M. M., De Rosa, C. T., Groten, J. P., Feron, V. J., Hansen, H., and Durkin, P. R. 1998. Estimation of toxicity of chemical mixtures through modeling of chemical interactions. *Environ. Health Perspect.* 106:1353-1360.
- Mumtaz, M. M., DeRosa, C. T., and Durkin, P. R. 1994. Approaches and challenges in risk assessments of chemical mixtures. In: *Toxicology of Chemical Mixtures: Case Studies, Mechanisms and Novel Approaches*. R. S. H. Yang, ed. San Diego: Academic Press. Ch. 22. pp. 565-597.
- Mumtaz, M. M., Durkin, P., Diamond, G. L., and Hertzberg, R. 1996. Exercises in the use of weight-of-evidence approach for chemical-mixture interactions. *J. Clean Technol. Environ. Toxicol. Occ. Med.* 5:339-345.
- Mumtaz, M. M., Poirier, K. A., and Colman, J. T. 1997. Risk assessment for chemical mixtures: Fine-tuning the hazard index approach. *J. Clean Technol. Environ. Toxicol. Occ. Med.* 6:189-204.
- Niederlehner, B. R., Cairns, J., and Smith, E. P. 1998. Modeling acute and chronic toxicity of nonpolar narcotic chemicals and mixtures to *Ceriodaphnia dubia*. *Ecotoxicol. Environ. Saf.* 39:136-146.
- NORA Mixed Exposures Team. 2004. *Mixed Exposures Research Agenda: A Report by the NORA Mixed Exposures Team*. U.S. Department of Health and Human Services, National Institute for Occupational Safety and Health, National Occupational Research Agenda, Cincinnati, OH. DHHA (NIOSH) Publication No. 2005-106.
- Norberg-King, T. J., Mount, D. I., Amato, J. R., Jensen, D. A., and Thompson, J. A. 1992. Toxicity identification evaluation: Characterization of chronically toxic effluents, Phase I. U.S. Environmental Protection Agency, Office of Research and Development, Duluth, MN. EPA/600/6-91/005F.
- Pohl, H. R., Hansen, H., and Chou, C. H. 1997. Public health guidance values for chemical mixtures: current practice and future directions. *Regul. Toxicol. Pharmacol.* 26:322-329.
- Pohl, H. R., Roney, N., Fay, M., Chou, C. H. S. J., Wilbur, S., and Holler, J. 1999. Site-specific consultation for a chemical mixture. *Toxicol. Ind. Health.* 15:470-479.
- Pohl, H. R., Roney, N., Wilbur, S., Hansen, H., and De Rosa, C. T. 2003. Six interaction profiles for simple mixtures. *Chemosphere.* 53:183-197.
- Price, B., Borgert, C. J., Wells, C. S., and Simon, G. S. 2002. Assessing toxicity of mixtures: The search for economical study designs. *Human. Ecol. Risk Assess.* 8:305-326.
- Putzrath, R. M. 1997. Estimating relative potency for receptor-mediated toxicity: reevaluating the toxicity equivalence factor (TEF) model. *Regul. Toxicol. Pharmacol.* 25:68-78.
- Putzrath, R. M. 2000. Reducing uncertainty of risk estimates for mixtures of chemicals within regulatory constraints. *Regul. Toxicol. Pharmacol.* 31:44-52.
- Rao, V. R. and Unger, A. 1995. Development of a risk assessment model for mixtures of dioxin congeners based on their competitive-binding to the Ah-receptor. *Environ. Carcino. Ecotox. Rev. C.* 13:53-74.
- Raucy, J. L. 1995. Risk assessment: Toxicity from chemical exposure resulting from enhanced expression of CYP2E1. *Toxicology.* 105:217-223.
- Reed, D. A. and Sterner, T. R. 2004. Using fractions to assess risk at petroleum release sites. Air Force Research Laboratory, Wright-Patterson AFB, OH. AFRL-HE-WP-TR-2004-0126.

- Reed, D. A., Sterner, T. R., Sarson, J. A. Total Petroleum Hydrocarbon Criteria Working Group (TPHCWG) field demonstration report: IRP Site 4, POL Area, Springfield ANG Base, Springfield, OH. Air Force Research Laboratory, Wright-Patterson AFB, OH. AFRL-HE-WP-TR-2003-0099.
- Reeves, W. R., Barhoumi, R., Burghardt, R. C., Lemke, S. L., Mayura, K., McDonald, T. J., Phillips, T. D., and Donnelly, K. C. 2001. Evaluation of methods for predicting the toxicity of polycyclic aromatic hydrocarbon mixtures. *Environ. Science Technol.* 35:1630-1636.
- Robinson, P. and MacDonell, M. 2004. Priorities for mixtures health effects research. *Environ. Toxicol. Pharmacol.* 18:201-213.
- Ross, K. E. and Bidwell, J. R. 2003. Assessing the application of an additive model to estimate toxicity of a complex effluent. *J. Environ. Qual.* 32:1677-1683.
- Safe, S. H. 1994. Polychlorinated biphenyls (PCBs): environmental impact, biochemical and toxic responses, and implications for risk assessment. *Crit. Rev. Toxicol.* 24:87-149.
- Sanderson, J. T. and van den Berg, M. 1999. Toxic equivalency factors (TEFs) and their use in ecological risk assessment: A successful method when used appropriately. *Human. Ecol. Risk Assess.* 5:43-52.
- Schoeny, R. S. and Margosches, E. 1989. Evaluating comparative potencies: developing approaches to risk assessment of chemical mixtures. *Toxicol. Ind. Health.* 5:825-837.
- Schoof, R. A. 2003. Guide for incorporating bioavailability adjustments into human health and ecological risk assessments at U.S. Department of Defense facilities. Part 1: Overview of metals bioavailability. Army Environmental Center, Aberdeen Proving Ground, MD.
- Schroder, J. L., Basta, N. T., Payton, M., Willson, J. A., Carlson, R. I., Janz, D. M., and Lochmiller, R. L. 2003. Ecotoxicological risks associated with land treatment of petrochemical wastes. I. Residual soil contamination and bioaccumulation by cotton rats (*Sigmodon hispidus*). *J. Toxicol. Environ. Health A.* 66:305-325.
- Schulte, P. 2003. Challenges for risk assessors. *Human. Ecol. Risk Assess.* 9:439-445.
- Schwartz, P. F., Gennings, C., and Chinchilli, V. M. 1995. Threshold models for combination data from reproductive and developmental experiments. *J. Am. Statist. Assoc.* 90:862.
- Seed, J., Brown, R. P., Olin, S. S., and Foran, J. A. 1995. Chemical-mixtures - Current risk assessment methodologies and future-directions. *Regul. Toxicol. Pharmacol.* 22:76-94.
- Sevigny, J. H., Tindal, M. J., Robins, G. L., Staudt, W., and Serbin, L. 2003. Importance of different volatile petroleum hydrocarbon fractions in human health risk assessment. *Human. Ecol. Risk Assess.* 9:987-1001.
- Sexton, K., Beck, B. D., Bingham, E., Brain, J. D., DeMarini, D. M., Hertzberg, R. C., O'Flaherty, E. J., and Pounds, J. G. 1995. Chemical mixtures from a public health perspective: The importance of research for informed decision making. *Toxicology.* 105:429-441.
- Simini, M., Checkai, R. T., and Maly, M. E. 2000. Tri-Service remedial project manager's handbook for ecological risk assessment. U.S. Army Environmental Center, Aberdeen Proving Ground, MD. SFIM-AEC-ER-CR-200015.
- Simmons, J. E. 1995. Chemical mixtures: Challenge for toxicology and risk assessment. *Toxicology.* 105:111-119.
- Simmons, J. E., Richardson, S. D., Speth, T. F., Miltner, R. J., Rice, G., Schenck, K. M., Hunter, E. S., and Teuschler, L. K. 2002. Development of a research strategy for integrated technology-based toxicological and chemical evaluation of complex mixtures of drinking water disinfection byproducts. *Environ. Health Perspect.* 110:1013-1024.
- Smolders, R., Bervoets, L., Wepener, V., and Blust, R. 2003. A conceptual framework for using mussels as biomonitors in whole effluent toxicity. *Human. Ecol. Risk Assess.* 9:741-760.
- Spongler, R. F. and Falk, H. 2002. Future directions of environmental public health research. ATSDR's 2002-2010 agenda for six priority focus areas. *Int. J. Hyg. Environ. Health.* 205:77-83.

- Sterner, T. R., Robinson, P. J., Mattie, D. R., and Burton, G. A. 2004. Analysis of algorithms predicting blood:air and tissue:blood partition coefficients from solvent partition coefficients for use in complex mixture physiologically based pharmacokinetic/pharmacodynamic modeling. Air Force Research Laboratory, Wright-Patterson AFB, OH. AFRL-HE-WP-TR-2004-0032.
- Suter, G. W. 2004. Bottom-up and top-down integration of human and ecological risk assessment. *J. Toxicol. Environ. Health A*. 67:779-790.
- Swartz, R. C. 1999. Consensus sediment quality guidelines for polycyclic aromatic hydrocarbon mixtures. *Environ. Toxicol. Chem.* 18:780-787.
- Tardif, R., Charest-Tardif, G., Brodeur, J., and Krishnan, K. 1997. Physiologically based pharmacokinetic modeling of a ternary mixture of alkyl benzenes in rats and humans. *Toxicol. Appl. Pharmacol.* 144:120-134.
- Teuschler, L. K. and Hertzberg, R. C. 1995. Current and future risk assessment guidelines, policy, and methods development for chemical mixtures. *Toxicology*. 105:137-144.
- Teuschler, L. K. and Simmons, J. E. 2003. Approaching DBP toxicity as a mixture problem. *J. Am. Water Works Assoc.* 95:131-138.
- Teuschler, L. K., Gennings, C., Stiteler, W. M., Hertzberg, R. C., Colman, J. T., Thiyagarajah, A., Lipscomb, J. C., Hartley, W. R., and Simmons, J. E. 2000. A multiple-purpose design approach to the evaluation of risks from mixtures of disinfection by-products. *Drug. Chem. Toxicol.* 23:307-321.
- Teuschler, L. K., Groten, J. P., Hertzberg, R. C., Mumtaz, M. M., and Rice, G. 2001. Environmental chemical mixtures risk assessment: current approaches and emerging issues. *Comments Toxicol.* 7:453-493.
- Teuschler, L., Klaunig, J., Carney, E., Chambers, J., Conolly, R., Gennings, C., Giesy, J., Hertzberg, R., Klaassen, C., Kodell, R., Paustenbach, D., and Yang, R. 2002. Support of science-based decisions concerning the evaluation of the toxicology of mixtures: a new beginning. *Regul. Toxicol. Pharmacol.* 36:34-39.
- Thomas, R. S., Rank, D. R., Penn, S. G., Zastrow, G. M., Hayes, K. R., Yu, T., Pande, K., Lewis, M., Jovanovich, S. B., and Bradfield, C. A. 2002. Application of genomics to toxicology research. *Environ. Health Perspect.* 110:919-923.
- Tilson, H. A. 2000. Neurotoxicology risk assessment guidelines: developmental neurotoxicology. *Neurotoxicology*. 21:189-194.
- Travis, C. C., Bishop, W. E., and Clarke, D. P. 2003. The genomic revolution: What does it mean for human and ecological risk assessment? *Ecotoxicology*. 12:489-495.
- Twerdock, L. E., Burton, D. T., Gardner, H. S., Shedd, T. R., and Wolfe, M. J. 1997. The use of nontraditional assays in an integrated environmental assessment of contaminated ground water. *Environ. Toxicol. Chem.* 16:1816-1820.
- Twerdok, L. E. 1999. Development of toxicity criteria for petroleum hydrocarbon fractions in the Petroleum Hydrocarbon Criteria Working Group approach for risk-based management of total petroleum hydrocarbons in soil. *Drug Chem. Toxicol.* 22:275-291.
- U.S. EPA. 1986. Guidelines for the Health Risk Assessment of Chemical Mixtures. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, D.C. EPA/630/R-98/002.
- U.S. EPA. 1989. Risk assessment guidance for Superfund. Volume I: Human health evaluation manual (Part A). Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, D.C. EPA/540/1-89/002, PB90-155581.
- U.S. EPA. 1997a. Ecological risk assessment guidance for Superfund: Process for designing and conducting ecological risk assessments. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, DC. EPA 540-R-97-006.
- U.S. EPA. 1997b. Priorities for ecological protection: An initial list and discussion document for EPA. U.S. Environmental Protection Agency, Washington, DC. EPA/600/S-97/002.

- U.S. EPA. 1998. Guidelines for ecological risk assessment. U.S. Environmental Protection Agency, Washington, DC. EPA/630/R-95/002F.
- U.S. EPA. 2000. Supplementary guidance for conducting health risk assessment of chemical mixtures. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC. EPA/630/R-00/002.
- U.S. EPA. 2002a. Guidelines Establishing Test Procedures for the Analysis of Pollutants; Whole Effluent Toxicity Test Methods; Final Rule. Federal Register, Washington, D.C. 40 CFR Part 136.
- U.S. EPA. 2002b. Framework for cumulative risk assessment. External review draft. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, D.C. EPA/630/P-02/001A.
- U.S. EPA. 2002c. Guidance on cumulative risk assessment of pesticide chemicals that have a common mechanism of toxicity. Office of Pesticide Programs, U.S. Environmental Protection Agency, Washington, D.C.
- U.S. EPA. 2002d. PPRTV derivation support document for total petroleum hydrocarbons. U.S. Environmental Protection Agency, National Center for Environmental Assessment, Superfund Health Risk Technical Support Center, Cincinnati, OH. SRC SF 01-031/10-16-2002.
- U.S. EPA. 2003. Human health toxicity values in Superfund risk assessments. Memorandum from Michael B. Cook, Director, Office of Superfund Remediation and Technology Innovation. Memorandum to Superfund National Policy Managers, Regions 1 - 10. Dated 5 Dec 2003. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, D.C. OSWER Directive 9285.7-53.
- U.S. EPA. 2004a. Integrated Risk Information System (IRIS). U.S. Environmental Protection Agency, Washington, D.C. Updated 22 Dec 2003. www.epa.gov/iris/.
- U.S. EPA. 2004b. An examination of EPA risk assessment principles and practices. Office of the Science Advisor, U.S. Environmental Protection Agency, Washington, D.C. EPA/100/B-04/001.
- USACHPPM. 2000. Standard practice for wildlife toxicity reference values. U.S. Army Center for Health Promotion and Preventative Medicine, Aberdeen Proving Ground, MD. Technical Guide No. 254.
- van den Berg, M., Peterson, R. E., and Schrenk, D. 2000. Human risk assessment and TEFs. Food Addit. Contam. 17:347-358.
- van den Brink, P. J., Hartgers, E. M., Gylstra, R., Bransen, F., and Brock, T. C. M. 2002. Effects of a mixture of two insecticides in freshwater microcosms: II. Responses of plankton and ecological risk assessment. Ecotoxicology. 11:181-197.
- Van Leeuwen, C. J., Verhaar, H. J. M., and Hermens, J. L. M. 1996. Quality criteria and risk assessment for mixtures of chemicals in the aquatic environment. Human. Ecol. Risk. Assess. 2:419-425.
- Verhaar, H. J. M., Morroni, J. R., Reardon, K. F., Hays, S. M., Gaver, D. P., Carpenter, R. L., and Yang, R. S. H. 1997. A proposed approach to study the toxicology of complex mixtures of petroleum products: the integrated use of QSAR, lumping analysis and PBPK/PD modeling. Environ. Health Perspect. 105:179-195.
- Viau, C. 2002. Biological monitoring of exposure to mixtures. Toxicol. Lett. 134:9-16.
- Vighi, M. and Calamari, D. 1996. Quality objectives for aquatic life: The problem of mixtures of chemical substances. Human. Ecol. Risk. Assess. 2:412-418.
- Villa, S., Vighi, M., Casini, S., and Focardi, S. 2003. Pesticide risk assessment in a lagoon ecosystem. Part II: effect assessment and risk characterization. Environ. Toxicol. Chem. 22:936-942.
- Vorhees, D. J. and Butler, O. L. 1999. Calculation of human health risk-based screening levels (RBSLs) for petroleum. Drug. Chem. Toxicol. 22:293-310.

- Vorhees, D. J., Weisman, W. H., and Gustafson, J. B. 1999. Total Petroleum Hydrocarbon Criteria Working Group. Volume 5: Human health risk-based evaluation of petroleum release sites: Implementing the Working Group Approach. Amherst, MA: Amherst Scientific Publishers.
- Wentsel, R. S., La Point, T. W., Simini, M., Checkai, R. T., Ludwig, D., and Brewer, L. 1994. Procedural guidelines for ecological risk assessments at U.S. Army sites. Volume 1. Edgewood Research, Development and Engineering Center, Aberdeen Proving Ground, MD. ERDEC-TR-221.
- Wentsel, R. S., La Point, T. W., Simini, M., Checkai, R. T., Ludwig, D., and Brewer, L. W. 1996. Tri-Service procedural guidelines for ecological risk assessments. Volume 1. U.S. Army Edgewood Research, Development and Engineering Center, Aberdeen Proving Ground, MD.
- Wireman, J. R., Hooper, M. J., and Porter, R. C. 2003. Ecological risk assessment within the U.S. Department of Defense. *Human. Ecol. Risk Assess.* 9:1561-1579.
- Wireman, Jody R. 2004. Personal communication (e-mail) with Jody R. Wireman, Ph.D., U.S. Air Force Institute for Environment, Safety and Occupational Health Risk Analysis, Brooks City Base, TX. 14 Feb 2004.
- Yang, R. S. H. 1998. Some critical issues and concerns related to research advances on toxicology of chemical mixtures. *Environ. Health Perspect.* 106:1059-1063.
- Yang, R. S., Thomas, R. S., Gustafson, D. L., Campaign, J., Benjamin, S. A., Verhaar, H. J., and Mumtaz, M. M. 1998. Approaches to developing alternative and predictive toxicology based on PBPK/PD and QSAR modeling. *Environ. Health Perspect.* 106:1385-1393.
- Zeliger, H. I. 2003. Toxic effects of chemical mixtures. *Arch. Environ. Health.* 58:23-29.