2. OPTIONS AND ISSUES FOR THE ASSESSMENT OF JOINT TOXIC ACTION OF CHEMICAL MIXTURES

In general, mixtures health or risk assessment focuses on three methods: use of data for the actual mixture of concern (also called the whole mixture), use of data for a similar mixture, or use of data on the components of the mixture. These methods are listed in order of preference.

2.1. MIXTURE OF CONCERN (WHOLE MIXTURE, ORIGINAL MIXTURE)

When exposure data and health effects data are available for the mixture of concern, use of this data has traditionally been the preferred approach (EPA 1986). Data on the mixture of concern are rarely available. When available, such data tend to be for complex mixtures that are considered a health hazard because they are generated in large quantities and are thought to cause adverse health effects. In addition, the exposures of concern generally occur at the source of the mixture. An example is coke oven emissions and fires. A mathematical model (N-Gas) has been developed to examine the toxic interactions of combustion gases consisting of six and seven gases to predict smoke toxicity from fires (Levin 1996; Levin et al. 1987a, 1987b, 1991, 1995). Health effects data are also available on pesticides, many of which are mixtures, often of isomers or congeners along with degradation products. A series of studies initiated by the National Institute of Environmental Health Sciences (NIEHS), however, focused on a mixture of 25 groundwater contaminants commonly associated with hazardous waste sites (Yang 1994). A similar study conducted on pesticide and fertilizer contaminants reported some evidence of cytogenetic damage (Kligerman et al. 1993; Yang 1994). These studies might be suitable as the basis for a mixture MRL or health guideline value, but this potential application of the studies does not appear to have been investigated.

The advantage of using data on the mixture of concern is that any interactions among the components of the mixture should be represented by the health effects data for the whole mixture. Limitations of the use of whole mixture data include the uncertainties regarding the extent to which the mixture from the exposure assessment “matches” the mixture that is the basis for the health criterion, due to changes in mixture composition with time and distance from the release, and/or differences in the original mixture. Thus, for most exposure scenarios, the mixture of concern will likely not be identical to the mixture that is the basis for the health criterion, even when it is called by the same name (e.g., toxaphene, PCBs). Further guidance on this topic is provided in Step 2 in Sections 4.2.1 and 4.2.3 of this guidance, and in the ATSDR (2001) interaction profile guidance.

2.2. SIMILAR MIXTURE

If no adequate data are available on the mixture of concern, but health effects data or guidance values are available on a similar mixture, the risk or health assessment may be based on the health effects data for the similar mixture, if the mixtures are sufficiently similar (EPA 1986, 2000). Sufficiently similar mixtures are those having the same chemicals but in slightly different proportions, or having most but not all chemicals in common and in highly similar proportions. In addition, the mixtures and their components have similar fate, transport, and health effects, whereas insufficiently similar mixtures may not. For example, JP-5 from different sources is considered similar because it is produced to meet uniform specifications, and differences from one source to another are thought to be minor (ATSDR 1998a). Gasoline from different sources was not considered sufficiently similar because of the wide range of formulations (ATSDR 1995a; Pohl et al. 1997). In addition, gasoline is a relatively heterogeneous mixture whose components have widely differing fate and transport characteristics (ATSDR 1995a, 1999). Consequently, receptor populations are likely to be exposed to subsets of the original components, and the subsets (or fractions) are not sufficiently similar to the original mixture (see Section 1.1).

Another method that has been used for risk assessment of similar mixtures is the comparative potency method. In this procedure, data for a set of similar mixtures are used to estimate a scaling factor that relates cancer potency derived from a chronic animal study or human epidemiology study to potency in a simpler assay, such as a mouse skin painting study. Then the cancer potency factor for an additional similar mixture for which only data from the simpler assay are available can be estimated using this scaling factor (Calabrese 1991; EPA 2000; Hertzberg et al. 1999; NRC 1988). This procedure has been used in the estimation of human cancer risk from combustion emissions from various sources (Albert et al. 1983; Lewtas 1985, 1988). Methods for noncarcinogenic effects are beginning to be developed (EPA 2000).

2.3. COMPONENTS

Due to the lack of suitable health criteria for the mixture of concern or a similar mixture, approaches involving the components of a mixture are commonly used for the incidental mixtures associated with hazardous waste sites. These methods are based on an assumption that the exposures or the responses to the mixture components are additive.

Dose Addition, also known as concentration addition, simple similar action, and similar joint action, assumes that the components of a mixture behave as concentrations or dilutions of one another, differing

only in their potencies (Bliss 1939; Finney 1971). The dose-response curves are parallel (i.e., the regression lines of probits on log doses are parallel), and tolerance (or susceptibility) to the components is completely positively correlated (the organisms most susceptible to chemical A also will be most susceptible to chemical B). The response to the mixture can be predicted by summing the doses of the components after adjusting for the differences in potencies. Dose addition is considered most appropriate for mixtures with components that affect the same endpoint by the same mode of action (EPA 1986, 1990, 2000). It has been suggested that the requirement for parallel dose-response curves and complete correlation of tolerances may be too stringent (e.g., Plackett and Hewlett 1952; Svendsgaard and Hertzberg 1994), and that in the low-dose region in which the response is linear, dose additivity may hold for independently acting chemicals as well (Svendsgaard and Hertzberg 1994). Dose addition is the underlying assumption of the hazard index method and the toxic equivalency factor (TEF) approach (Sections 2.3.1 and 2.3.3).

Response Addition, also known as simple independent action and independent joint action (Bliss 1939), assumes that the chemicals act independently and by different modes of action. Tolerance (or susceptibility) to one component may or may not be correlated with tolerance to another. The organisms most susceptible to chemical A may also be most susceptible to chemical B (complete positive correlation) or may be least susceptible to chemical B (complete negative correlation), or the susceptibilities to the two chemicals may be statistically independent. The response to the mixture can be predicted from the responses to the components and the correlation of tolerances. Response addition is the underlying assumption of an approach to cancer risk assessment for mixtures and ACGIH's approach to assessing the hazard of occupational exposure to agents that act independently (Sections 2.3.5 and 3.1).

Additional detail regarding dose and response addition is provided in Appendix A.

Evidence to Support the Use of Dose Additivity Models

Acute studies using overtly toxic doses of binary mixtures have shown that deviations from dose additivity generally are not remarkable in mammals (e.g., Smyth et al. 1969, 1970; Withey and Hall 1975). Toxicity studies on guppies and frogs using mixtures of 3 to as many as 50 components also tend to indicate that deviations from dose addition are not substantial (e.g., Dawson 1994; Hermens et al. 1985; Konemann 1981). Deviations from dose additivity were generally less than a factor of five. A number of investigations have focused on the low dose (low response) area. In a series of 4-week feeding studies by the TNO Nutrition and Food Research Institute, mixtures of four chemicals were administered orally to rats at doses of the individual chemicals below the no-observed-adverse-effect level (NOAEL), equivalent to the NOAEL, and at an adverse effect level. These studies gave results for renal toxicity that

were consistent with dose additivity or that appeared less-than-dose-additive. The mixtures consisted of four similarly acting nephrotoxicants (Feron et al. 1995) and four dissimilarly acting nephrotoxicants (Jonker et al. 1993). The above conclusions are based partly on the investigators' observations, and partly on a reanalysis of the individual animal data using exponential dose-response functions, performed for ATSDR (Mumtaz et al. 1998). Results of other studies by the same institute on mixtures of eight (Jonker et al. 1990) and nine (Groten et al. 1997) dissimilarly acting chemicals reported few effects when the doses of the individual components of the mixture were subtoxic.

Other studies, however, indicate co-exposure to subthreshold doses or environmental doses of chemicals that affect the same target organs (though not by the same mechanism) can result in adverse effects. An acute study of a mixture of subthreshold doses of 1,1,1-trichloroethane, trichloroethylene, and tetrachloroethylene in rats resulted in adverse effects on the liver (Stacey 1989). Although cadmium and lead affect the hematological system through different mechanisms, dietary exposures of rats to these metals at doses that did not significantly affect hemoglobin and hematocrit when given individually, resulted in significant decreases in hemoglobin and hematocrit when given as a mixture (Mahaffey and Fowler 1977; Mahaffey et al. 1981). A series of studies initiated by the NIEHS on a mixture of 25 groundwater contaminants from hazardous waste sites indicated that toxic effects can result from long-term exposure to mixtures in which each of the components is present at doses expected to be subtoxic (Yang 1994). A similar NIEHS study conducted on pesticide and fertilizer contaminants reported some evidence of cytogenetic damage (Kligerman et al. 1993; Yang 1994). The individual components were not tested in the NIEHS studies, so further analysis of the data for interactions or additivity is problematic, but the authors used doses of the individual chemicals that were expected to be without effect. Epidemiological studies of children have indicated that lead and arsenic, and lead and cadmium, may interact at environmental levels of exposure to produce adverse neurobehavioral consequences in children (Marlowe et al. 1985; Moon et al. 1985). On the other hand, some studies in animals and humans (e.g., Berman et al. 1992; Caprino et al. 1983; Drott et al. 1993; Harris et al. 1984) have reported apparent thresholds for interactions (see also Section 2.3.7).

Data that indicate a lack of interactive effects may not, however, mean there is no interaction. The biological systems currently in use may not be sensitive enough to detect interactions, and the power of many joint toxic action studies may be insufficient to conclusively demonstrate additivity or interactions. Newer techniques, such as genomics and proteomics, may provide tools for detecting toxicological interactions at very low dose levels.

Based on the above evidence and concerns, the dose-additivity assumption may be a reasonable *default* assumption for chemicals with similar effects or the same target organ in the low dose range. Use of the

dose-additivity assumption is likely to produce estimates of health hazard that range from appropriate to somewhat conservative, and which are therefore protective of public health.

2.3.1. Hazard Index

The hazard index approach uses the assumption of dose additivity to assess the noncancer health effects of a mixture from the data on the components. EPA has adopted the term "hazard index" for this approach, which appears to have originated in 1972 (see Section 3.5). The approach is used or recommended by a number of agencies (ACGIH 2000; EPA 1986, 1989a; Mumtaz et al. 1994a, 1997; National Academy of Sciences [NAS] 1974; National Research Council [NRC] 1989; OSHA 1993, 2001). Exposures or doses for the various components of the mixture are scaled by a defined level of exposure generally regarded as "acceptable" or "safe" by the agency performing the assessment. The defined levels could be ATSDR MRLs, EPA reference doses (RfDs) or reference concentrations (RfCs), ACGIH threshold limit values (TLVs), or OSHA permissible exposure limits (PELs). The general equation for the hazard index (*HI*) is:

$$(a) \quad HI = \frac{E_1}{DL_1} + \frac{E_2}{DL_2} + \cdots + \frac{E_n}{DL_n} \qquad or \qquad (b) \quad HI = \sum_{i=1}^{n} \frac{E_i}{DL_i} \tag{1}$$

In equation 1(a), E_1 is the level of exposure to the first chemical in the mixture and DL_1 is some defined level of exposure to the first chemical, E_2 and DL_2 are the corresponding levels for chemical 2, and the summation can extend to any number of chemicals, signified by the *n*. Equation 1(b) simply expresses the same idea more succinctly, where *i* is the i^{th} chemical. Each chemical-specific ratio (e.g., E_1 / DL_1) is called a hazard quotient (HQ). Therefore, the hazard index can be expressed as the sum of the hazard quotients:

$$HI = \sum_{i=1}^{n} HQ_i \tag{2}$$

When the hazard quotient for a single chemical exceeds unity, concern for the potential hazard of the chemical increases. Similarly, when the hazard index for a mixture exceeds unity, concern for the potential hazard of the mixture increases.

Separate hazard indexes are estimated for each pathway and exposure duration of concern. For a given duration, hazard indexes are summed across pathways that affect the same receptor population.

The obvious advantage of this method is its simplicity. Because it is based on the assumption of dose additivity, the hazard index method is most appropriately applied to components that cause the same effect by the same mechanism of action. In practice, it may be applied to components with different target organs (sometimes as a screening measure). The method is frequently applied to components with

the same critical target organ or critical effect (effect that is the basis for the MRL, RfD, or other health guideline), without regard to mechanism of action. For Superfund risk assessments, strong evidence is required to indicate that two compounds producing adverse effects on the same organ system, although by different mechanisms, should not be treated as dose additive (EPA 1989a). See also the discussion in Section 2.3 (*Evidence to Support the Use of Dose Additivity Models*).

The hazard index method does not take into account interactions among the components of the mixture.

Additional information on this method is provided in EPA (1986, 1989a).

2.3.2. Target-organ Toxicity Dose Modification to Hazard Index Method

The target-organ toxicity dose (TTD) method, which is a refinement of the hazard index method, was devised in order to accommodate the assessment of mixtures whose components do not all have the same critical effect. In addition, it takes into account the reality that most components of waste-site-related mixtures affect other target organs at doses higher than those that cause the critical effect. These other effects may vary from component to component and may be important in assessing the health effects of the mixture. EPA (1989a) suggested that separate hazard indexes be estimated for all endpoints of concern. EPA further suggested that the RfD be used not only in generating hazard quotients for the critical effect of a component, but also in estimating hazard quotients for effects that occur at higher exposure levels. As acknowledged by EPA (1989a) and demonstrated by Mumtaz et al. (1994a, 1997), this practice may overestimate the hazard for effects occurring at exposure levels higher than those associated with the critical effect. The use of TTDs was therefore suggested (Mumtaz and Colman 1992; Mumtaz et al. 1997), and is consistent with the recommendations of EPA (1986) and NRC (1989), discussed in Sections 3.4 and 3.5. TTDs are developed for the chemicals that affect an endpoint at a dose higher than that for the critical effect for the same chemical. A TTD for each endpoint of concern is calculated using appropriate MRL (or RfD) methodology, and then used in estimating the endpoint-specific hazard indexes. The MRL (or RfD) is used for the critical effect for each chemical and the TTD is used for the other endpoints of concern for the chemical. When any of the endpoint-specific hazard indexes exceeds unity, concern for the potential hazard of the mixture increases.

The derivation of TTDs for use in assessment of the joint toxic action of chemical mixtures is analogous to the derivation of MRLs, and should follow the applicable portions of ATSDR (1996a) MRL guidance. TTDs are based on the other major characteristic effects of a chemical, which are known to occur at the same or higher exposure levels as the critical effects. Like the derivation of an MRL, the derivation of a TTD is not recommended for an endpoint that is affected only at the relatively high levels of exposure

associated with severe effects. Because the purpose of TTD derivation is to support the estimation of endpoint-specific hazard indexes (Guidance Manual Section 2.3.1; Mumtaz et al. 1994a, 1997), TTD derivations should be performed for endpoints that are common to more than one component of a given mixture. In addition, endpoints identified as concerns in populations exposed to the mixture should be considered.

In common with MRLs, TTDs are specific for route and exposure period. The TTD should be based on the highest NOAEL that does not exceed a lowest-observed-adverse-effect level (LOAEL) for the particular endpoint, as determined from the information in toxicological profiles, including the Levels of Significant Exposure Tables. If such a NOAEL is not available, the TTD would be based on the lowest LOAEL for that endpoint. Additional considerations, as for MRL derivation, are that the NOAEL or LOAEL used as the basis for the TTD should be from a representative, quality study, for the same route and exposure period as the TTD. When data for the exposure duration of concern are not available, a TTD derived for one duration may sometimes be applicable for other duration(s) of the same route, if supported by the overall database. An additional uncertainty factor may be applied to extrapolate across exposure durations, based on scientific judgment. Dose adjustments and interspecies, intraspecies, and LOAEL to NOAEL extrapolation should be performed as for an MRL. When suitable data are available, and when appropriate, TTDs can also be derived using benchmark dose (BMD) modeling (Crump 1984, 1995; EPA 2001; Gaylor et al. 1998) to define the BMD, which is used in place of a NOAEL as the basis for TTD derivation, similar to the procedure for MRL derivation.

For example, suppose that chemicals 1, 2, 3, and 4 are commonly found in combination in completed exposure pathways involving intermediate oral exposure. The intermediate oral MRLs for chemicals 1 and 2 are based on hepatic effects, and for chemicals 3 and 4 are based on renal effects and developmental effects, respectively. Each of these endpoints also is affected by at least one other mixture component for which it is not the critical effect. Other major effects in common for two or more of these chemicals for this route and duration include neurological and reproductive effects. In addition, chemical 1 causes immunological effects and chemical 4 causes endocrine (adrenal) effects during intermediate oral exposure. At levels of exposure that cause high mortality, chemical 1 also causes hematological effects in rats. This information is summarized in Table 4.

Table 4. Endpoints Affected by Chemicals 1, 2, 3, and 4

	AFFECTED BY			
ENDPOINT	Chemical 1	Chemical 2	Chemical 3	Chemical 4
Hematological	with mortality	No	No	No
Hepatic	Yes—MRL	Yes—MRL	No	Yes
Renal	Yes	No	Yes—MRL	Yes
Endocrine (adrenal)	No	No	No	Yes
Immunological	Yes	No	No	No
Neurological	Yes	Yes	Yes	No
Developmental	Yes	Yes	Yes	Yes—MRL

The endpoints of concern chosen for TTD derivation, based on the critical effects of the chemicals and on other major effects in common for this set of chemicals, are hepatic, renal, neurological, and developmental effects. These endpoints are shown in bold italicized print in the table. Since adrenal and immunological effects each are caused by only one chemical, and are not the critical effects for any of the components of the mixture, the estimation of endpoint-specific hazard indexes is not needed for these endpoints, and TTDs are accordingly not developed. For a different mixture of chemicals that included chemical 1, the immunological endpoint may warrant TTD derivation if at least one other chemical in the mixture also causes this effect. Similar reasoning would apply for chemical 4 and adrenal effects. The hematological effects are not a suitable basis for TTD derivation for chemical 1 not only because they are caused by only one chemical, but also because they occurred only at levels of exposure that caused significant mortality.

For the purposes of illustration, a TTD for renal effects will be derived for chemical 1. The intermediate oral MRL for chemical 1 is 0.15 mg/kg/day based on a NOAEL of 15 mg/kg/day for hepatic effects in experimental animals given the chemical orally for an intermediate duration. The NOAEL was divided by an uncertainty factor of 100 (10 for interspecies and 10 for intraspecies variability) to estimate the MRL. The LOAEL for hepatic effects in the same study was 30 mg/kg/day. The NOAEL and LOAEL for renal effects in the this study were 30 and 45 mg/kg/day, and were the most reliable data for this effect. In addition, the NOAEL was the highest NOAEL for this effect. A TTD_{RENAL} of 0.3 mg/kg/day is derived by dividing the $NOAEL_{RENAL}$ of 30 mg/kg/day by an uncertainty factor of 100 (10 for interspecies and 10 for intraspecies variability). Derivation of TTDs for the other effects would proceed in a similar manner.

Following derivation of the TTDs, endpoint-specific hazard indexes are calculated as follows:

$$
\begin{aligned}
(a) \quad HI_{HEPATIC} &= \frac{E_1}{MRL_1} + \frac{E_2}{MRL_2} + \frac{E_4}{TTD_{4\,HEPATIC}} \\
(b) \quad HI_{RENAL} &= \frac{E_1}{TTD_{1\,RENAL}} + \frac{E_3}{MRL_3} + \frac{E_4}{TTD_{4\,RENAL}} \\
(c) \quad HI_{NEURO} &= \frac{E_1}{TTD_{1\,NEURO}} + \frac{E_2}{TTD_{2\,NEURO}} + \frac{E_3}{TTD_{3\,NEURO}} \\
(d) \quad HI_{DEV} &= \frac{E_1}{TTD_{1\,DEV}} + \frac{E_2}{TTD_{2\,DEV}} + \frac{E_3}{TTD_{3\,DEV}} + \frac{E_4}{MRL_4}
\end{aligned}
\qquad \textbf{(3)}
$$

where $HI_{ENDPOINT}$ is the hazard index for indicated endpoint (*HEPATIC*, *RENAL*, *NEURO* [neurological], *DEV* [developmental]), E_i is the exposure for the i^{th} chemical (1, 2, 3, or 4 in the above example), MRL_i is the MRL for the i^{th} chemical, and TTD_i is the TTD for the i^{th} chemical for the indicated endpoint. (If an MRL is not available, a suitable RfD can be used.) Although developmental toxicity is the critical effect for only one of the four chemicals, all four produce the effect, and it is conceivable that it may be a sensitive effect for the mixture. Neurological effects are not the critical effect for any of the chemicals, but three of the chemicals cause this effect at equivalent or higher exposure levels than associated with the critical effect. Thus, use of the TTD modification of the hazard index for mixtures of chemicals that do not have the same critical effect may increase the understanding of the potential impact of the mixture on public health. Additional information regarding this method is provided by Mumtaz et al. (1994a, 1997).

The development of TTDs is analytically intensive. TTDs have been developed for a variety of chemicals in a pilot study (Mumtaz et al. 1997) and are being developed in ATSDR interaction profiles. The derivations in the interaction profiles are subjected to a review process that is similar to that for MRLs. The development of these values for all substances that are currently the subjects of toxicological profiles, for each duration and route of exposure, would be problematic. To address the issue of practicality, the method could be limited to those situations where clarification of the public health hazard is needed (as described in Sections 4.2.2 and 4.2.3), the TTD effort could be focused on chemicals that frequently occur together in mixtures of concern needing such clarification of public health hazard, and TTD determinations could be made available to health assessors through an easily accessible and readily updated medium, such as the ATSDR website, or through interaction profiles. If the method proves useful, the addition of TTDs to the toxicological profiles could be considered, to be phased in as new profiles are developed and existing profiles are updated. The TTDs could be developed and reviewed in conjunction with MRLs.

2.3.3. Weight-of-Evidence Modification to the Hazard Index

As noted above, the hazard index method does not incorporate information on interactions among components of the mixture. A weight-of-evidence (WOE) method proposed by Mumtaz and Durkin (1992) was the first systematic attempt to address this need. The method implemented and expanded on the suggestion made by the NRC (1989) that, in recognition of the difficulties of quantifying interactions, an uncertainty factor (UF) be used to account for interactions among components of a mixture (Section 3.5). The method was designed to modify the hazard index to account for interactions, using the weight of evidence for interactions among pairs of mixture components. Although subsequent experience with the algorithm used to generate the interactions hazard index has revealed that it does not handle changes in proportions of mixture components in a reasonable manner, the method is useful qualitatively for predicting whether hazard may be greater or less than indicated by the hazard index.

The method evaluates data relevant to joint action for each possible pair of chemicals in the mixture in order to make qualitative binary weight-of-evidence (BINWOE) determinations for the effect of each chemical on the toxicity of every other chemical. Two BINWOEs are needed for each pair: one for the effect of chemical A on the toxicity of chemical B, and another for the effect of chemical B on the toxicity of chemical A. The BINWOE determination is a classification that indicates the expected direction of an interaction (greater than additive, less than additive, additive, or indeterminate), and scores the data qualitatively, using an alphanumeric scheme that takes into account mechanistic understanding, toxicological significance, and relevance of the exposure duration, sequence, bioassay (*in vitro* versus *in vivo*), and route of exposure. The alphanumeric terms in the classification scheme can then be converted to a single numerical score, by multiplying the corresponding direction factor by the data quality weighting factor. Although the earlier publications of the WOE method did not discuss the need for BINWOE determinations to take into account target organ (Durkin 1995; Mumtaz and Durkin 1992), experience in application of the WOE method, including preparation of the ATSDR interaction profiles and a study by Mumtaz et al. (1998), has indicated that the WOE evaluations should be target-organ specific.

The qualitative BINWOE classifications are shown in the left column of Table 5 and the direction factors and data quality weighting factors are shown in the far right column. An alphanumeric (qualitative) BINWOE classification of >II.B.2.a.i for the effect of one chemical on the toxicity of another thus corresponds to greater-than-additive interaction, mechanistic data on related chemicals, inferred toxicological significance, different duration or sequence, *in vivo* data, and anticipated route of exposure. The corresponding BINWOE score is +1(0.71)(0.71)(0.79)(1)(1)=+0.40.

The weight of evidence method used the numerical BINWOE scores as the interaction terms in an equation that took into account the doses and potencies (through use of hazard quotients) of the components of the mixture, and calculated a composite score for interactions, WOE_N, that was intended to be an expression of the strength of the evidence that interactions may be toxicologically significant relative to the highest possible level of certainty that would be possible for the particular mixture. Details are provided in Appendix B. The WOE_N was used to modify an interactions uncertainty factor (UF_I), as follows:

$$HI_I = HI_{add} \times UF_I^{WOE_N} \quad \textbf{(4)}$$

where HI_I is the interactions-adjusted hazard index and HI_{add} is the hazard index based on additivity. An uncertainty factor of 10 was chosen in exercises illustrating the method (Mumtaz and Durkin 1992; Mumtaz et al. 1994a). Because this algorithm does not does not handle changes in proportions of mixture components in a reasonable manner, a qualitative WOE method is used, as described in the following paragraph.

Table 5. Binary Weight-of-Evidence Scheme for the Assessment of Chemical Interactions*

Classification		Factor
Direction of Interaction		**Direction**
=	Additive	0
>	Greater than additive	+1
<	Less than additive	–1
?	Indeterminate	0
Quality of the Data		**Weighting**
Mechanistic Understanding		
I.	Direct and Unambiguous Mechanistic Data: The mechanism(s) by which the interactions could occur has been well characterized and leads to an unambiguous interpretation of the direction of the interaction.	1.0
II.	Mechanistic Data on Related Compounds: The mechanism(s) by which the interactions could occur have not been well characterized for the chemicals of concern but structure-activity relationships, either quantitative or informal, can be used to infer the likely mechanisms(s) and the direction of the interaction.	0.71
III.	Inadequate or Ambiguous Mechanistic Data: The mechanism(s) by which the interactions could occur has not been well characterized or information on the mechanism(s) does not clearly indicate the direction that the interaction will have.	0.32
Toxicological Significance		
A.	The toxicological significance of the interaction has been directly demonstrated.	1.0
B.	The toxicological significance of the interaction can be inferred or has been demonstrated for related chemicals.	0.71
C.	The toxicological significance of the interaction is unclear.	0.32
Modifiers		
1.	Anticipated exposure duration and sequence.	1.0
2.	Different exposure duration or sequence.	0.79
a.	*In vivo* data	1.0
b.	*In vitro* data	0.79
i.	Anticipated route of exposure	1.0
ii.	Different route of exposure	0.79

Weighting Factor = Product of Weighting Scores: Maximum = 1.0, Minimum = 0.05

BINWOE = Direction Factor x Weighting Factor: Ranges from –1 through 0 to +1

*Adapted from Mumtaz and Durkin (1992) and Mumtaz et al. (1994a)

A qualitative WOE approach, focusing on application of the BINWOE scores to hazardous waste-site assessment, was suggested by Mumtaz and Durkin (1992). This approach is appropriate for a mixture where the scaled doses (hazard quotients) for all of the components are similar, or toxicologically significant. The qualitative BINWOE scores for the components, if similar in direction, are the basis for a conclusion. For example, consider a mixture of four components, all present at toxicologically significant levels. The number of possible chemical pairs in a mixture of N components is $(N^2-N)/2$. Thus, this mixture of 4 components has 6 pairs of components and potentially 11 BINWOEs. Suppose nine of the BINWOEs are greater-than-additive (positive) with alphanumeric classifications indicating a relatively high degree of confidence, and the remaining three BINWOEs are additive (0), also with relatively high degrees of confidence. In this case, the weight of evidence suggests that the mixture is likely to pose a greater hazard than that indicated by the hazard index.

A likely pattern of qualitative BINWOEs for a mixture is a mixed pattern (some greater than additive, some less than additive, and some additive BINWOEs). In this case, the qualitative WOE approach is extended to include conversion of the qualitative BINWOE scores to numerical scores, and summing the scores to give a combined score. If the combined BINWOE score is positive and significantly different from zero, then the weight of evidence suggests that the mixture is likely to pose a greater hazard than indicated by the hazard index. Conversely, if the combined BINWOE score is negative and significantly different from zero, then the weight of evidence suggests that the health hazard is unlikely to be greater than indicated by the hazard index. Professional judgment is used in the interpretation of the impact of the WOE on the hazard index.

Although the WOE method was developed for assessing interactions for noncarcinogenic effects, the qualitative WOE method is equally applicable to assessing interactions for carcinogenic effects.

The WOE method (Mumtaz and Durkin 1992; Mumtaz et al. 1994a) has undergone evaluation, and appeared to perform well qualitatively, and quantitatively under some circumstances. The application of the method for deriving BINWOE classifications was considered consistent by expert toxicologists who reviewed the results of exercises in which several teams of toxicologists and risk assessors independently determined BINWOE classifications for the same pairs of chemicals (Mumtaz et al. 1994b). In tests of the WOE method to predict the toxicity of some simple chemical mixtures to animals, BINWOEs for three pairs of chemicals qualitatively predicted whether the results of animal studies would be less-than-additive, additive, or greater-than-additive (Mumtaz et al. 1998). Used with an exponential dose-response model and dose addition to model relative kidney weights, the quantitative WOE method closely predicted the observed dose-response in female rats for intermediate-duration oral exposure to a mixture of four nephrotoxic chemicals with similar modes of action (Mumtaz et al. 1998). The observed dose-

response was less than dose additive. The BINWOEs were focused on renal toxicity, and the uncertainty factor used in the algorithm was 10. The WOE method underestimated the relative liver weights in the same animals. The observed dose-response for relative liver weight was slightly greater than dose additive. Thus, the WOE method did not predict toxicity to a target organ that was different from the one for which the BINWOEs were derived. The WOE method slightly overpredicted the observed dose-response for relative kidney weight in male rats for a mixture of dissimilarly acting nephrotoxins (in female rats, the data variability was so great that the exponential model did not fit the *observed* responses) (Mumtaz et al. 1998). Although these results are suggestive, limitations of this test of the complete WOE method include the substantial variability in the responses of individual animals, small numbers of animals per group, testing of only two dose levels of the mixtures, and lack of rationale for using relative organ weight as an index of toxicity (several other indicators of renal and hepatic toxicity were monitored in the studies that provided the experimental data [Jonker et al. 1993, 1996]).

A modification of the original WOE method was proposed by Eastern Research Group (ERG) and Durkin (1995) and has been further developed by and adopted as part of its mixtures guidance (EPA 2000). This modification includes a slightly different classification scheme and a different method of calculating the interactions-modified hazard index. The method encourages greater use of quantitative interaction data through the use of magnitude-of-interaction factors for each chemical pair. The classification scheme, while more integrated in nature, requires more judgment, and the type of quantitative interaction data required to estimate the magnitude factor is rarely available. The algorithm for this modification appears to handle changes in proportions of mixture components more reasonably than does the original algorithm, but additional evaluation with regard to predicting experimental results is desirable.

A basic assumption of both WOE methods is that interactive interference will not be significant. For example, if chemicals A and B interact in a certain way, the presence of chemical C will not cause the interaction to be substantially different. Thus, the assumption is that pairwise interactions will dominate in the mixture and adequately represent all the interactions.

Additional detail regarding both methods is provided in Appendix B, and detailed guidance for deriving BINWOE determinations and evaluating joint toxic action studies is presented in ATSDR (2001).

2.3.4. Toxic Equivalency and Relative Potency

The toxic equivalency and relative potency approaches also use the assumption of dose additivity to assess the health effects of a mixture. These approaches have been applied to mixtures that consist of a

class of chemicals, and are used when health effects information for one component of the mixture is sufficient to derive health criteria but for the other components of the mixture is less complete.

The toxic equivalency approach has been used with the CDDs and structurally related chemical classes such as the chlorinated dibenzo-*p*-furans (CDFs) and the coplanar PCBs (Ahlborg et al. 1994; ATSDR 1998b; EPA 1989b, 1994; Safe 1998; Van den Berg et al. 1998). This method estimates toxic equivalency factors (TEFs) for the various congeners in the mixture based on the key assumption that certain congeners exert effects such as carcinogenicity through a common receptor-mediated mechanism (Ah receptor), and therefore act in a dose additive manner. The TEF approach compares the potency of individual congeners, based on *in vitro* or acute *in vivo* data, with that of 2,3,7,8-TCDD, the best-studied of this chemical class. 2,3,7,8-TCDD is assigned a TEF of unity; the other TEFs (or relative potencies) are usually less than one. The concentrations or doses of each active congener are multiplied by their TEF values and then summed to give the total toxic equivalents (TEQs) of a mixture:

$$TEQs = \sum_{i=1}^{n} C_i \times TEF_i \qquad \textbf{(5)}$$

where C_i is the concentration (or dose) and TEF_i is the TEF for the i^{th} component of the mixture. The TEQ thus represents the concentrations of all the components as an equivalent concentration of the index chemical, 2,3,7,8-TCDD. The hazard or risk of exposure to the mixture is estimated by comparing the TEQs with MRLs or other health-based criteria (ATSDR environmental media evaluation guide [EMEG]; ATSDR screening, evaluation, and action levels) based on 2,3,7,8-TCDD (ATSDR 1998b; De Rosa et al. 1997a, 1998; Mumtaz and Hertzberg 1993; Pohl et al. 1995) or multiplying the TEQ (in appropriate units of mg/kg/day or mg/m^3) by a cancer slope factor or unit risk for 2,3,7,8-TCDD (EPA 1994, 1996; Mumtaz and Hertzberg 1993).

This approach is considered suitable for the assessment of health effects of dioxin-like compounds that are mediated through the Ah receptor, but is not applicable for those that are not (ATSDR 1998b). Carcinogenicity (at least in part), immunotoxicity, and developmental and reproductive toxicity (the basis for oral MRLs) are thought to be mediated through the Ah receptor (ATSDR 1998b). Limitations to this method are that some of the nondioxin-like PCB congeners have been shown to inhibit or enhance responses to 2,3,7,8-TCDD, depending on dose and assay system (Birnbaum and DeVito 1995; Pohl and Holler 1995; Safe 1998); the range of TEF values estimated for some PCB congeners is very broad (Safe 1998); and a slope factor for 2,3,7,8-TCDD is not available on the Integrated Risk Information System (IRIS). The TEF approach continues to evolve and undergo additional testing and validation. ATSDR considers the approach less suitable for PCBs, and has derived MRLs for PCBs (ATSDR 2000). ATSDR is using the TEF method as a tool for assessing health effects of dioxin and dioxin-like compounds (primarily CDDs and CDFs) in soil (ATSDR 1998b; De Rosa et al. 1997a, 1998).

A similar approach, called a relative potency approach, has been developed for PAHs that have been classified as B2 carcinogens by EPA (ATSDR 1995b; EPA 1993). The relative potency factors are estimated on the basis of potency relative to that of benzo[a]pyrene in skin painting carcinogenesis studies. Benzo[a]pyrene is the best-studied of this class and has a cancer potency factor available on IRIS. The mechanistic underpinnings of the relative potency approach for the PAHs are less good, in terms of the additivity assumption. Some of the same issues as noted for the application of the TEF approach also are issues for the use of the relative potency method for PAHs, including nonadditive interactions among the PAHs.

2.3.5. Total Cancer Risk

A response addition approach has been recommended for the assessment of risk from mixtures of carcinogenic chemicals (De Rosa et al. 1993; EPA 1986, 2000; Mumtaz et al. 1994a; NRC 1989). The most conservative form of response addition, completely negative correlation of tolerances (i.e., individuals most sensitive to chemical A are least sensitive to chemical B and vice versa; see Appendix A) was recommended by EPA (1986). Accordingly, the response or risk for the mixture is the sum of the risks for the components:

$$Risk = \sum_{i=1}^{n} Risk_i = \sum_{i=1}^{n} d_i B_i \qquad \textbf{(6)}$$

where $Risk_i$ is the risk, d_i is the dose, and B_i is a potency parameter (slope factor or unit risk) for the i^{th} carcinogen. The equation is appropriate when risks for the individual chemicals are less than 0.01 and the sum of the individual risks is less than 0.1 (EPA 1989a). This equation is equivalent to dose addition if the dose-response curves for the chemicals are within the linear (low-dose) range, and have no threshold (EPA 1986, 2000). EPA (2000) recommends the response addition model for independent action (as in equation 18 of Appendix A) for cancer risk, noting that when component risks are small, the formula collapses to the simple addition of component risks (equation 6 above). Use of the IRIS values for slope factor or unit risk result in plausible upper bounds to the lifetime excess cancer risk of the components. Concern has been raised that summing upper bound risks may lead to unreasonably high estimates of the mixture risk, but an analysis by Kodell and Chen (1994) suggests that the error in the simple sum of the upper bound risks is small relative to other uncertainties, and Cogliano (1997) concluded that the sum of the upper bound risks provides useful information regarding the overall risk from mixtures of carcinogens.

2.3.6. The Integral Search System (ISS) for Ranking Hazards of Mixtures of Carcinogens

The ISS method (Woo et al. 1994), like the WOE method, uses data for binary mixtures to predict the hazard of exposure to mixtures of three or more chemicals. The method is carried out by a software package. The ISS integrates three EPA and National Cancer Institute databases on binary interactions of carcinogens with carcinogens, promoters, and inhibitors. It contains approximately 1,000 chemicals of 60 structural and functional classes. The ISS calculates a weighting ratio reflecting the ratio of greater-than-additive to less-than-additive interactions for the components of a mixture. The estimation of the weighting ratio is based on the interactions data for the chemical pairs in the mixture and, for those pairs lacking interactions data, on interactions between other members of the chemical classes to which the chemicals belong. The weighting ratio also incorporates judgments as to the relative effectiveness of the four type of interactions (synergism, promotion, antagonism, and inhibition) in modifying the hazard. Weighting ratios greater than unity indicate that the combined effect of the mixture components is expected to be greater-than-additive, whereas ratios less than unity indicate that the combined effect is expected to be less-than-additive.

In addition, ISS can be used to estimate a "concern level," which is based on the "inherent hazard" (the sum of the slope factors for the components, converted to an exponent index value), multiplied by the weighting ratio. The resulting score is converted back to a weighted total slope factor and to a corresponding concern level, ranging from low to high.

A serious limitation, however, is that ISS does not include exposure concentration or dose as part of this procedure. Another serious limitation is that the class-class interaction ratings for pairs of chemicals with no data tend to dominate the score. The attractive features of the ISS are that it calculates the weighting ratios automatically, it is applicable to mixtures with relatively large numbers of components, and it can accommodate the assessment of chemicals that are not presently included in the database as long as the chemical can be assigned to an appropriate class of chemicals within the database. Additional detail regarding ISS is provided in Appendix C.

2.3.7. PBPK, PBPK/PD, and Quantitative Structure Activity Relationships (QSAR)

PBPK and PBPK/PD techniques are beginning to be applied to problems in mixtures toxicology. For mixtures of two chemicals, PBPK and PBPK/PD models for the individual chemical are linked at the assumed point of interaction, frequently the hepatic metabolism term. Following validation of the assumed mechanism by comparing model predictions with experimental data, the model can be used to predict effects of co-exposure for different exposure scenarios. For example, binary PBPK models have

been developed to extrapolate from high-exposure inhalation studies of interactions of toluene and xylene in rats to low exposure in humans by the same route (Tardif et al. 1995) and to identify functional interaction thresholds for the joint toxicity of trichloroethylene and 1,1-dichloroethylene in the rat (El-Masri et al. 1996a). PBPK/PD models have been applied to further assess apparent interaction thresholds for the joint toxicity of trichloroethylene and 1,1-dichloroethylene (El-Masri et al. 1996b) and of kepone and carbon tetrachloride (El-Masri et al. 1996c) in the rat, and to extrapolate from high-dose studies of interactions of toluene and dichloromethane in animals to lower-dose exposures by a different route in humans (Pelekis and Krishnan 1997). As an example of the direct applicability to the assessment of potential hazard to human health, the study of toluene and dichloromethane illustrates the use of PBPK/PD modeling to estimate the effect of co-exposure to toluene on the induction of carboxyhemoglobinemia (adverse effect) by dichloromethane in humans at defined levels of exposure.

The above models deal with binary mixtures. Approaches to modeling simple mixtures of three or more components also are under development (Haddad and Krishnan 1998; Haddad et al. 1999a, 1999b; Krishnan and Pelekis 1995; Tardif et al. 1997). As with the models for binary mixtures, these models for three or more components are constructed by linking the models for the individual chemicals based on pairwise interaction mechanisms, and the model predictions are validated with experimental data. The reported predictions of the models may be directly useful in assessing the potential hazard of joint toxic action of the simple mixtures studied. For example, separate and linked PBPK models were used to estimate biological hazard indexes (based on blood concentrations of parent compound) for varying exposures and proportions of a three-chemical mixture (toluene, ethylbenzene, and m-xylene) (Haddad et al. 1999b). These biological hazard indexes may be relevant to the central nervous system effects of the compounds, which are considered to be due to the parent compounds.

A PBPK model for the BTEXs in the rat demonstrated the utility of this approach for predicting the blood concentrations of the parent compounds in rats following inhalation exposure to the mixture (Haddad et al. 1999a). Blood levels of the parent compounds may be relevant to central nervous system effects. The study further demonstrates that models linked on the basis of binary interactions adequately predict the inhalation pharmacokinetics of a four-component mixture.

An approach to dealing with complex mixtures is to model portions of the mixture as a single component or “lump.” This approach has been used to predict whether the metabolism of benzene to genotoxic metabolites is affected by the other components of gasoline in the mouse (Bond et al. 1998). A similar approach has been proposed and partially developed for studying the acute toxicology of JP-5, a Navy jet fuel that contains a complex mixture of petroleum hydrocarbons in the C9–C18 range (Verhaar et al. 1997; Yang et al. 1998). The focus is on the prediction of kinetics of JP-5 components in relevant tissues

after acute inhalation exposure and the resultant narcosis from the dissolution of hydrocarbons in the membranes of nerve cells. The approach involves the lumping of similar mixture components into a pseudocomponent, for which necessary chemical parameters such as tissue partition coefficients are estimated. QSARs are used to estimate necessary model parameters for pseudocomponents, such as tissue-blood and air-blood partition coefficients, and metabolic rate constants.

The binary, simple, and complex mixture models discussed above are being developed and validated with acute exposure data. Results of a study using PBPK modeling and experimental data obtained at intervals during a 2-year inhalation study on dichloromethane suggest that age of the animal and continuing exposure to this chemical produce changes in disposition and metabolism, such that the use of models based on acute data may not adequately predict intermediate and chronic exposure (Thomas et al. 1996). Exposure levels in this study were relatively high (2,000 ppm, 6 hours/day, 5 days/week) and, therefore, may or may not be applicable to low exposure. To date, few of these modeling efforts include extrapolation to humans. A PBPK/PD approach for carcinogenesis is under development, but has not yet been applied to mixtures or to extrapolate to humans (Yang et al. 1998).

PBPK and PBPK/PD models are being used to efficiently design experiments to test hypotheses of interaction mechanisms and to predict whether interactions may occur at low levels of exposure, so that testing can focus on mixtures of greater concern. As this field of research progresses, however, these models are expected to become useful in more direct assessment of potential hazard to human health (Haddad and Krishnan 1998). Examples of this direct application were provided previously in this section. PBPK and PBPK/PD models could be used to explore exposure scenarios involving different intakes, proportions, and routes of exposure for the mixture components (Haddad and Krishnan 1998). In addition, such models may be used as the basis for deriving health guideline values for the mixture of concern: PBPK/PD modeling may provide estimates of an "interaction threshold" (e.g., LED_{05}, lower 95% confidence limit on an effective dose associated with a 5% extra risk) for a simple mixture that could be used as a benchmark dose for derivation of a guidance value (Yang et al. 1998). Integration of PBPK/PD models with other approaches such as Monte Carlo simulation, response surface methodology, and QSAR is expected to further enhance predictive capability (El-Masri et al. 1997; Yang et al. 1998). Thus, there is a clear need for research that would provide the data that would allow the modeling to be predictive and effective.