5.0 FETAX TEST METHOD DATA AND RESULTS

5.1 Availability of Detailed FETAX Protocol

A comprehensive ASTM guideline for FETAX was published in 1991 and a revised guideline was published in 1998. The 1991 and 1998 versions of the ASTM FETAX Guideline are provided in **Appendix 10** and **11**, respectively. The protocol used in the FETAX Phase I Validation Study followed the 1991 ASTM Guideline (Bantle et al., 1994a). This guideline, with minor modifications, was followed in the Phase II (Bantle et al., 1994b), Phase III.1 (Bantle et al., 1996), Phase III.2 (Fort et al., 1998), and Phase III.3 (Bantle et al., 1999) Validation Studies. **Section 7.2** discusses each validation study and any protocol modifications. Unless noted otherwise, the 1991 ASTM FETAX Guideline was followed in the other FETAX studies.

5.2 Availability of Original and Derived FETAX Data

Original and derived data were obtained for each of the five FETAX validation studies from Dr. Bantle, the lead investigator.

5.3 Statistical Approach used to Evaluate FETAX Data

The statistical and non-statistical methods used by the individual investigator to analyze FETAX data obtained in their laboratory are described in **Section 2.1.13**. To obtain a consensus call for each substance tested in each validation study, the validation study management team determined the average of the calculated LC_{50}, EC_{50}, TI, and MCIG values among all replicate definitive tests (generally three replicate definitive tests per compound per participating laboratory). The conclusion as to the potential teratogenicity of a test substance was then based on the average TI and the average ratio of the MCIG to the LC_{50}. This method for achieving a consensus conclusion does not take into account the variability among laboratories in reaching their own conclusion as to the potential teratogenicity of the test substance. In contrast, NICEATM used a weight-of-evidence approach based on the results obtained for each laboratory. In this approach, a test substance was classified as positive in FETAX if a majority of laboratories obtained a positive result. Similarly, a test substance was classified as negative in

FETAX if a majority of laboratories obtained a negative result. In situations where an equal number of positive and negative studies were available for consideration, the test substance was classified as equivocal and excluded from any analysis.

5.4 FETAX Test Results for Individual Substances

FETAX test data from 276 separate studies involving 137 individual substances (not including environmental samples) were located, reviewed, extracted, and entered into the NICEATM FETAX database (**Appendix 2** contains substances tested without metabolic activation, **Appendix 3** contains substances tested with metabolic activation). Sources for these data included peer-reviewed literature (including studies accepted for publication) and non peer-reviewed book chapters. Information provided in abstracts and manuscripts not accepted for publication were not considered. All 137 substances had been tested using FETAX without metabolic activation; 35 of these 137 substances had also been tested with metabolic activation.

A summary of the responses for substances tested multiple times, as well as the weight-of-evidence conclusion, are provided in **Appendix 6**.

FETAX test results are classified in the database as positive or negative based on the criteria provided in the ASTM FETAX Guideline (1991, 1998) (i.e., positive if the TI value is greater than 1.5 or if the MCIG/LC_{50} ratio is less than 0.30). Also, in keeping with a recent study (Fort et al., 2000a), FETAX test results are classified as positive if the TI value is greater than 3.0. In addition, consistent with recent studies where both the TI value and the MCIG/LC_{50} ratio were considered together in classifying FETAX results, compounds are classified as positive based on obtaining concordant positive results for both endpoints using a TI value greater than 1.5 and an MCIG/LC_{50} ratio less than 0.30 (Bantle et al., 1999), or using a TI value greater than 3.0 and an MCIG/LC_{50} ratio less than 0.30 (Fort et al., 2000a), negative if neither the TI value or the MCIG/LC_{50} ratio were positive, and equivocal if only one of the two endpoints were positive. Due to the lack of quantitative *X. laevis* malformation data in the majority of publications, this endpoint was not considered in the assessment of performance characteristics by NICEATM. The importance of using agent-specific characteristic abnormalities in classifying materials as positive in FETAX is discussed in **Section 6.6.2**.

5.4.1 FETAX Test Results Without Metabolic Activation

Of the 137 substances tested without metabolic activation, 105 substances were tested only once. The remaining 32 substances were tested in multiple studies. The number of multiple studies ranged from three to 14. TI data were available for all studies, while MCIG data were available for 96 (70%) of the test substances. Qualitative data on malformations observed in *X. laevis* without metabolic activation were available for 35 substances, including three environmental samples. Quantitative malformation data by test substance concentration were not provided in any study.

5.4.2 FETAX Test Results With Metabolic Activation

Of the 35 substances tested with metabolic activation, 21 were tested only once. The remaining 14 substances were tested in multiple studies ranging from three and eight. TI data were available for all studies, while MCIG data were available for 27 (77%) of the test substances. Qualitative data on malformations observed in *X. laevis* without metabolic activation were available for six substances; quantitative malformation data by test substance concentration were not provided in any study.

5.5 FETAX Test Results With Binary Mixtures

FETAX has been used also to assess the teratogenicity and embryotoxicity of binary mixtures, in the absence of metabolic activation only. The rates of malformation by binary mixtures are expected to depend on the mode of teratogenesis for the component substances of the mixture. For those mixtures comprised of substances that follow the same modes of action, concentration-addition rates of malformation are expected. In contrast, response-addition rates are expected for those mixtures containing substances with different modes of action.

Dawson and Wilke (1991a, b) tested a total of 12 defined binary mixtures (**Table 3**) using FETAX. In the first study (Dawson and Wilke, 1991a), three mixtures were tested using ratios of 0:1, 3:1, 1:1, 1:3, and 1:0. Compound selection was based on their different modes of teratogenicity and their mortality/malformation index (MMI). All of the mixtures tested

Table 3. Teratogenicity/Embryolethality of Binary Mixtures Tested in FETAX (Dawson and Wilke, 1991a, b).

Mixture	Mixture Ratio	Toxic Units[1] (mixture)	Displayed Effect	Malformations Observed
Semicarbazide:Isoniazid	3:1	1.09	Concentration addition	Skeletal
	1:1	1.03		
	1:3	1.02		
Valproic acid:Pentanoic acid	3:1	1.02	Concentration addition	Head and osmoregulatory
	1:1	1.03		
	1:3	0.98		
Butyric acid:Pentanoic acid	3:1	0.98	Concentration addition	Head and osmoregulatory
	1:1	1.06		
	1:3	1.06		
Hydroxyurea:Isoniazid	3:1	1.15	Response addition	Skeletal, head, visceral and osmoregulatory
	1:1	1.29		
	1:3	1.29		
Isoniazid:6-Aminonicotinamide	3:1	1.23	Response addition	Skeletal and eye
	1:1	1.27		
	1:3	1.15		
Isoniazid:Retinoic acid	3:1	1.23	Response addition	Skeletal and mouth
	1:1	1.22		
	1:3	1.30		
Hydroxyurea:Retinoic acid	3:1	1.25	Response addition	Skeletal, head, mouth, visceral, and osmoregulatory
	1:1	1.39		
	1:3	1.35		
6-Aminonicotinamide: Retinoic acid	3:1	1.24	Response addition	Eye and mouth
	1:1	1.36		
	1:3	1.27		
Retinoic acid:Nicotine	3:1	1.35	No interaction[2]	Mouth
	1:1	1.76		
	1:3	1.37		
Isoniazid: ß-Aminopropionitrile	3:1	0.97	Response addition	Connective tissue lesions (typical of osteolathyrism), visceral edema, gut mis-coiling, facial malformations
	1:1	1.01		
	1:3	1.00		
Valproic acid:Butyric acid	3:1	1.01	Response addition	Reduced head size, visceral/cranial edema, poor gut coiling, skeletal kinking, occasional mouth/eye defects.
	1:1	0.96		
	1:3	0.98		
Isoniazid:Valproic acid	3:1	1.33	Response addition	Not provided
	1:1	1.53		
	1:3	1.19		
Semicarbazide:Isoniazid (embryolethality)	3:1	1.12	Response addition	Not evaluated
	1:1	1.12		
	1:3	1.11		
Hydroxyuera:Isoniazid (embryolethality)	3:1	1.52	Response addition[3]	Not evaluated
	1:1	1.35		
	1:3	1.15		

[1] Toxic unit = EC_{50} in mixture/EC_{50} alone; [2] Effect was not greater than that observed for each compound individually;
[3] One concentration (3:1) produced a TU value indicative of antagonism. This was attributed to an excess concentration of isoniazid, which effected the efficiency of the absorption of hydroxyurea in the mixture.

displayed response addition. In the second study (Dawson and Wilke, 1991b), nine binary mixtures of developmental toxicants were tested for teratogenicity. Each of the mixtures was tested using ratios of 0:1, 3:1, 1:1, 1:3, and 1:0. The mixtures were analyzed using the toxic unit (TU) method, which is based on an individual substance's EC_{50} value being defined as 1.0 TU for malformation induced in *X. laevis* by that substance (Dawson, 1991). Three of these mixtures had calculated TU values near 1.0, which is indicative of concentration addition. The remaining six mixtures displayed TU values greater than 1.0, indicative of response addition. The investigators concluded that the results of these studies indicated that a developmental endpoint could be useful in the assessment of joint toxic action studies (Dawson and Wilke, 1991a, b).

Dawson and Wilke (1991b) also tested two binary mixtures—semicarbazide:isoniazid and hydroxyurea:isoniazid—for lethal effects. The semicarbazide:isoniazid mixture displayed response addition, although both substances are known to have the same mode of action for teratogenic effects (osteolathyrism). This suggested to the investigators that the two substances followed different modes of action for embryolethality. The hydroxyurea:isoniazid mixture was also found to display response addition, although an antagonistic TU value for the 3:1 ratio was observed. This result was attributed to the high relative concentration of isoniazid reducing the efficiency of hydroxyurea absorption and, therefore, its contribution to the mixture's lethality (Dawson and Wilke, 1991b).

A mixture comprised of ten aliphatic carboxylic acids was tested using FETAX malformations as an endpoint (Dawson, 1991). The results of this study are shown in **Table 4**. Based on the TU method, Dawson concluded that the mixtures displayed a concentration additive response.

Dawson and Wilke (1996) conducted an extensive evaluation of malformation dose-response curves for binary mixtures of differently acting teratogenic substances in FETAX. This study was purported to be the first to examine substances in combination where only one of the agents was present at an effective dose in the mixture. The substances tested were 6-AN, β-aminoproprionitrile, benzoic hydrazide, butyric acid, cytarabine, 2-ethylhexanoic acid,

Table 4. Malformation Information for Ten Carboxylic Acid Mixtures Tested in FETAX (Dawson, 1991)

Concentration of Mixture (mL)[1]	Number of Embryos Exposed/ Survivors	Number of Malformations	Malformations Observed[2]
0	75/75	12	Skeletal kinking (3), microcephaly (2), gut coiling (2), eye edema/blister (2), general edema (2), mouth (1)
2	75/75	26	Microcephaly (6), gut coiling (5), skeletal kinking (5), general edema (4), mouth (3), eye edema/blister (3)
4	75/75	41	Microcephaly (18), gut coiling (14), mouth (4), eye edema/blister (4), skeletal kinking (1)
5	75/75	53	Microcephaly (25), gut coiling (20), eye edema/blister (3), mouth (2), skeletal kinking (2), general edema (1)
6	75/75	76	Microcephaly (35), gut coiling (32), skeletal kinking (3), eye edema/blister (3), general edema (2), mouth (1)
8	75/75	127	Microcephaly (56), gut coiling (52), eye edema/blister (8), skeletal kinking (5), mouth (4), general edema (2)

[1] Total solution in exposure dishes = 10 mL
[2] Numbers in parenthesis indicate the number of embryos that exhibit the preceding malformation.

5-fluorouracil, hydroxyurea, isoniazid, penicillamine, pentanoic acid, trans-retinoic acid, thiosemicarbazide, and valproic acid. The binary mixtures were prepared such that in mixtures of the agents was present in almost ineffective concentrations. For 16 pairs of substances, the 1:1 mixture was slightly more effective in inducing malformations than would be expected based on additivity alone. In contrast, the 1:3 and 3:1 mixtures were not more effective than the effective agent in that combination alone.

FETAX tests have also been performed on mixtures of mixed xylenes and toluene (Kononen and Gorski, 1997). The LC_{50} and EC_{50} values in these experiments were less than predicted, thus indicating possible synergism between the two substances. However, confidence levels were not calculated at the various concentration levels and therefore further testing would need to be performed to confirm this interpretation.

Based on the information presented, FETAX appears to be useful for conducting toxicity assessments on substance mixtures. Both embryolethality and malformation are relevant endpoints to be evaluated when assessing mixtures, although modes of action also need to be considered. Embryolethality is best used for non-teratogenic mixtures since the mode of action does not effect the outcome of testing. Teratogens are best evaluated using the malformation endpoint due to the likelihood of separate modes of action for malformation and lethality that would make interpretation of results difficult. The need for a developmental malformation endpoint was stressed as a means of identifying chronic toxicity rendered by developmental abnormalities (Dawson and Wilke, 1991b).

5.6 Use of Coded Chemicals and Compliance with GLP Guidelines

Coded substances were not used in the Phase I Validation Study (Bantle et al., 1994a), but were used in the Phase II (Bantle et al., 1994b), Phase III.1 (Bantle et al., 1996), Phase III.2 (Fort et al., 1998), and Phase III.3 (Bantle et al., 1999) Validation Studies. It does not appear that blind coding was used in any other FETAX study. However, in the Phase II Validation Study, the same preset test concentrations were used by all laboratories for each test substance.

FETAX validation studies were not conducted in compliance with national or international GLP guidelines, nor were they generally conducted at facilities at which GLP studies are normally conducted. It does not appear that any FETAX study was conducted in compliance with GLP guidelines.

5.7 Availability of Non-Audited FETAX Data

None of the FETAX data obtained by NICEATM had been audited by a Quality Assurance Unit. However, copies of all original data collected in the five FETAX validation studies were obtained by NICEATM for a possible independent audit. (see **Section 7.0**). Original data was not sought by NICEATM for any other FETAX study.

5.8 Section 5 Conclusions

A detailed ASTM FETAX protocol was first published in 1991. With minor exceptions, the FETAX validation studies followed this protocol. Original and derived data were obtained for all five FETAX validation studies only; no attempt was made by NICEATM to obtain any other original FETAX data.

The averaging method used in the FETAX validation studies for achieving a consensus call does not take into account the variability among laboratories in reaching their own conclusion as to the potential teratogenicity of the test substance. In contrast, NICEATM used a weight-of-evidence approach based on the results obtained for each laboratory. The relative appropriateness and merits of these two approaches should be evaluated.

The FETAX database includes 276 separate studies involving 137 substances. All 137 substances had been tested using FETAX without metabolic activation; 35 of these 137 substances had also been tested with metabolic activation. FETAX has been used to assess the teratogenicity and embryotoxicity of defined binary mixtures. Both embryolethality and malformation are relevant endpoints to be evaluated when assessing mixtures, although modes of action also need to be considered. NICEATM has concluded that the potential utility of FETAX for this purpose merits additional investigation.

Except for the most recent four of five FETAX validation studies, it does not appear that blind coding was used to eliminate potential bias in any other FETAX study. Also, it does not appear that any FETAX studies were conducted in compliance with national or intemational GLP guidelines.

The effect of these two issues on the quality of the data in the FETAX database is difficult to ascertain.