Appendix L

[This Page Intentionally Left Blank]

Appendix L-1

Outlier Characterization for the 3T3 and NHK NRU Test Methods with the RC Regression

[This Page Intentionally Left Blank]

L.1 Discordant Results for the 3T3 and NHK NRU Test Methods and RC Millimole Regression

The RC millimole regression and each *in vitro* NRU test method were used to identify discordant results among the reference chemicals tested in the validation study (i.e., those for which the rodent LD_{50} was not accurately predicted by the *in vitro* NRU IC_{50}). Discordant chemicals are also referred to as outliers. Once identified, discordant chemicals were then evaluated for common characteristics that may assist in determining the types of chemicals that are not suited for use in the 3T3 and NHK NRU test methods to determine starting doses for acute systemic toxicity assays. **Sections L.1.1** and **L.1.2** identify discordant chemicals for the RC weight regression and the combined 3T3 or the combined NHK NRU weight regressions (see **Tables L1-1 and L1-2**). Discordant chemicals for the millimole regressions are discussed in this appendix.

L.1.1 Identification of Discordant Chemicals

For each *in vitro* NRU test method, predicted LD_{50} values for the reference chemicals were determined using the geometric mean IC_{50} of the three geometric mean laboratory values in the RC millimole regression. Discordant chemicals were identified using the RC method (Halle 1998): a difference greater than 0.699 (or log 5) for log observed LD_{50} (in mmol/kg) minus log predicted LD_{50} identifies a chemical as discordant (i.e., an outlier) (see **Appendix J-1** for the 3T3 NRU test method and **Appendix J-3** for the NHK NRU test method for the predicted LD_{50} values for each chemical). **Table L1-1** lists the discordant chemicals for the RC millimole regression

Table L1-1 Discordant Chemicals for the 3T3 and NHK NRU Test Methods and the RC Millimole Regression

	3T3 NRU	NHK NRU
	Acetaminophen (+)	
	Arsenic III trioxide (-)	*Arsenic III trioxide (-)*
		Aminopterin (-)
5-Aminosalicylic acid		5-Aminosalicylic acid (+)
Busulfan	**Busulfan (-)**	**Busulfan (-)**
Caffeine		Caffeine (-)

Table L1-1 Discordant Chemicals for the 3T3 and NHK NRU Test Methods and the RC Millimole Regression

	3T3 NRU	NHK NRU
Cycloheximide	**Cycloheximide (-)**	**Cycloheximide (-)**
Dibutyl phthalate	***Dibutyl phthalate (+)***	***Dibutyl phthalate (+)***
	Dichlorvos (-)	Dichlorvos (-)
	Diethyl phthalate (+)	***Diethyl phthalate (+)***
Digoxin	***Digoxin (-)***	
Disulfoton	***Disulfoton (-)***	***Disulfoton (-)***
	Endosulfan (-)	Endosulfan (-)
Epinephrine bitartrate	Epinephrine bitartrate (-)	Epinephrine bitartrate (-)
Ethanol	**Ethanol (+)**	**Ethanol (+)**
	Fenpropathrin (-)	*Fenpropathrin (-)*
	Hexachlorophene (-)	
Lindane	*Lindane (-)*	*Lindane (-)*
Mercury II chloride	Mercury II chloride (-)	Mercury II chloride (-)
		Methanol (+)
Nicotine	Nicotine (-)	Nicotine (-)
Paraquat	Paraquat (-)	Paraquat (-)
Parathion	***Parathion (-)***	***Parathion (-)***
Phenobarbital	Phenobarbital (-)	Phenobarbital (-)
Phenylthiourea	Phenylthiourea (-)	Phenylthiourea (-)
	Physostigmine (-)	Physostigmine (-)
Potassium cyanide	Potassium cyanide (-)	Potassium cyanide (-)
Propylparaben	Propylparaben (+)	Propylparaben (+)
	Sodium hypochlorite (+)	Sodium hypochlorite (+)
		Sodium oxalate (-)
	Sodium selenate (-)	Sodium selenate (-)
	Strychnine (-)	*Strychnine (-)*
Thallium I sulfate	Thallium I sulfate (-)	
Triethylenemelamine	Triethylenemelamine (-)	Triethylenemelamine (-)
1,1,1-Trichloroethane		
Verapamil HCl	**Verapamil HCl (-)**	***Verapamil HCl (-)***

[1]Log LD_{50} (mmol/kg) = 0.435 log IC_{50} (mM) + 0.625. Log LD50 (mmol/kg) for discordant chemicals are > 0.699 from the RC regression.

(-) - toxicity underpredicted by the *in vitro* assay (i.e., the LD_{50} value predicted by the IC_{50} is higher than the *in vivo* LD_{50} value); (+) - toxicity overpredicted by the model (i.e., the LD_{50} value predicted by the IC_{50} is lower than the rodent reference LD_{50} value).

Bolded chemicals have active metabolites *in vivo* (see **Appendix F-2**).

Chemicals that showed evidence of insolubility (i.e., precipitates) during testing (see **Table 5-8**) are identified by italics.

Table L1-1 shows a comparison of the 22 RC outlier chemicals tested (identified in **Table 3-2**) with the outliers identified when using the 3T3 and NHK NRU results with the RC regression. Using the RC method of identifying discordant chemicals (Halle 1998), there were 30 discordant chemicals for the 3T3 assay and 31 discordant chemicals for the NHK assay. For the 58 RC chemicals that were tested, both test methods confirmed the outlier status of 19 of the 22 RC outliers tested, but the chemicals not identified as outliers were somewhat different for the 3T3 and NHK assays. The 3T3 NRU cytotoxicity test method did not identify caffeine, 5-aminosalicylic acid and 1,1,1-trichloroethane as outliers and the NHK test method did not identify digoxin, thallium sulfate, and 1,1,1-trichloroethane as outliers. The 3T3 assay identified four chemicals as outliers that were not identified as outliers by the RC: arsenic trioxide; hexachlorophene; acetaminophen; and diethyl phthalate. The NHK assay identified five chemicals as outliers that were not identified as outliers by the RC: aminopterin; arsenic trioxide; sodium oxalate; diethyl phthalate; and methanol.

L.1.2 Evaluation of Discordant Chemicals

To determine the attributes that may be used in the future to identify chemicals that would result in discordant predictions, a number of physico-chemical characteristics were evaluated for their frequency of occurrence among the discordant chemicals versus the entire set of test chemicals. This section provides a summary of these analyses for the discordant chemicals identified among the validation set based on the RC millimole regression and the RC outlier criteria.

Physical Characteristics

A number of physical characteristics were evaluated for their frequency of occurrence in the set of discordant chemicals versus the entire set of reference chemicals. The characteristics chosen were those that were assumed to be available, or relatively easy to measure, for new chemicals that may be tested in these assays to determine starting doses for acute systemic toxicity assays. The characteristics examined included chemical class, molecular weight, boiling point, IC_{50}, pH, and log K_{ow} (i.e., log octanol:water partition coefficient). Unfortunately, these attributes were not available for all chemicals. For example, log K_{ow} was available for only 52 of 72 (72%) chemicals and boiling point was available for only 26

of 72 (36%) chemicals. For boiling points > 200°C, 9 of 13 (69%) chemicals were outliers using the 3T3 NRU test method results and 8 of 13 (62%) chemicals were outliers using the NHK NRU test method results. For molecular weight > 400 g/mole, 5 of 7 (71%) chemicals were outliers using the 3T3 NRU test method results and 3 of 7 (43%) chemicals were outliers using the NHK NRU test method results. For log K_{ow} > 3, 9 of 12 (75%) chemicals were 3T3 outliers and 8 of 12 (67%) chemicals were NHK outliers.

Chemical Class

Examination of outliers by chemical class for the RC regression showed that 3 of 3 (100%) organophosphates were outliers in both *in vitro* NRU test methods.

Solubility

Another attribute that may cause a chemical to be discordant is the lack of solubility in the solvent or when added to the tissue culture media. Since the SMT expected the toxicity of insoluble chemicals to be underpredicted in the *in vitro* NRU cytotoxicity assays, chemical tests with precipitates were noted and compared with the discordant chemicals. However, insolubility was not consistently associated with the discordant chemicals for which toxicity was underpredicted. For example, only 5 of the 24 (21%) underpredicted chemicals identified by applying the 3T3 NRU test method results to the RC regression exhibited signs of insolubility in the 3T3 assay in at least one laboratory (see **Table 5-8** for chemicals that had precipitates in the NRU assays). Additionally, evidence of insolubility was noted for dibutyl phthalate and diethyl phthalate in both assays, but toxicity was overpredicted rather than underpredicted.

For the 3T3 NRU, 25 chemicals showed evidence of insolubility in the 3T3 assay in at least one lab. Eleven (4%) of these chemicals were outliers. For the NHK NRU, 24 chemicals showed evidence of insolubility in at least one laboratory. Nine (38%) of the 24 chemicals were outliers.

Metabolism

The SMT expected that the toxicity of chemicals metabolized *in vivo* to active compounds (see **Table 3-7**) would be underpredicted *in vitro* by 3T3 and NHK cells, which have little or no metabolic capability. Of the 72 reference chemicals used to validate the 3T3 and NHK NRU test methods, 22 (31%) chemicals are known to have active metabolites *in vivo* and only 10 (45%) of these chemicals were classified as outliers for the 3T3 NRU. Of these 10 chemicals, the toxicity of six (60%) chemicals was underpredicted, while the toxicity of four (40%) chemicals was overpredicted. For the NHK NRU, nine (29%) of the 31 outlier chemicals are metabolized to active metabolites. Four (44%) of nine are negative outliers (predicted to be less toxic than they are). See **Table 3-7** for chemicals with active metabolites. Thus, the fact that a chemical has active metabolites does not necessarily indicate that its toxicity will be underpredicted by the *in vitro* NRU cytotoxicity test methods.

Halle (1998) reported similar findings for the RC database (i.e., approximately half of the chemicals metabolized to active metabolites were outliers and half were not).

Mechanism of Action

The contribution of mechanism of action to discordant status was assessed by developing additional regressions for the prediction of LD_{50} values from the *in vitro* NRU IC_{50} values. Both the 3T3 and NHK NRU test method data and the RC database were used for this evaluation. The RC database was used because:

- it is the largest published compilation of IC_{50x} and rodent LD_{50} data
- the RC IC_{50x} data were highly correlated with the 3T3 and NHK NRU IC_{50} data collected for the 58 chemicals in common, and especially for the *in vitro* 3T3 NRU IC_{50} data (see **Figures 5-3** and **5-4**)

In evaluating the contribution of mechanism of action to discordant status based on the RC weight regression, chemicals with mouse LD_{50} data only were excluded from the comparison. RC rat data only were used because:

- rat and mouse data should not be combined, regardless of the high correlation of their LD_{50} data reported in the RC publication on their subset of rat and mouse

LD_{50} values (see **Section 4.1.2**), since rats and mice may not have the same sensitivity to individual chemicals

- the majority of data used in the RC regression were rat data (282 rat data points and 65 mouse data points) (Halle 1998)
- the great majority of acute oral systemic toxicity testing is performed with rats.

The linear regression developed from the 282 rat data points in the RC (see **Appendix K-4**) using weight units is shown in **Table 6-2** and **Figure 6-5**. **Table 6-2** shows that the RC rat weight regression was not significantly different from the weight regression for the complete RC database when slopes and intercepts were simultaneously compared (goodness of fit F test; p=0.961).

The following boxes characterize the 30 discordant chemicals (i.e., outliers) for the 3T3 NRU by counting the number of outliers in each category and comparing to the total number of chemicals in the category.

Physical Form	Number of Outliers/Total in Category
Solid	23 outliers/54 solids
Liquid	7 outliers/18 liquids

Boiling Point (in degrees C)	Outliers/Total
No info	14/34
< 100	1/8
100-200	1/5
200-300	3/4
300-400	5/6
408	1/1
960	0/1
1500	0/1
others decompose, sublime, or BPs are given @ < atmospheric pressure	

Molecular Weight (g/mol)	Outliers/Total in Class
< 150	3/21 chemicals (but no info on MeOH or CCl4)
> 150-200	6/14
200-300	13/20
300-400	3/11
400-500	3/4
500-600	1/1
600-700	0/1
700-800	1/1

Chemical class	Number of Outliers/Total in Class	
Alcohols	3/10	
Carboxylic acids	4/12	
Heterocyclic	7/14	
Mercury compounds	1/1	
Organophosphorous	3/3	(2 were organothiophosphorous cmpds)
Polycyclic	1/2	
Sulfur compound	1/2	
Sodium compound	2/5	
Organometallic	1/1	
Amide	1/3	
Amine	1/1	
Arsenical	1/2	
Boron compounds	0/1	
Cadmium compounds	0/1	
Cyclic hydrocarbon	1/2	
Halogenated hydrocarbon	1/3	
Hydrocarbon	1/2	
Ketone	0/1	
Lithium compound	0/1	
Metal compound	1/1	
Nitrile	0/1	
Phenol	0/1	
Potassium compound	1/2	
Chlorine compound	1/2	
Nitrogen compound	1/1	
Chromium compound	0/1	
Fluorine compound	0/1	
Oxygen compound	1/1	
Selenium compound	1/1	

3T3 IC50 (mM)	Outliers/Total
≤ 0.0001	0/2
0.0001 - 0.001	1/2
0.001 - 0.01	1/4
0.01 - 0.1	9/13
0.1 - 1	15/23
1-10	2/12
10-100	1/9
> 100	1/5

No IC50s for 2 chemicals

3T3 pH	Outliers/Total
< 7.6	0/9
7.6	0/0
7.7	1/1
7.8	0/1
7.9	3/6
8	5/12
8.1	11/18
8.2	3/6
8.3	3/8
8.4	1/5
8.5	0/1
> 8.5	3/4

log Kow	Outliers/Total
< -4	0/1
-4 to –1	2/7
-1 to 0	3/7
0 to 1	4/7
1 to 2	5/13
2 – 3	1/5
3 - 4	5/8
4 - 5	2/2
6 – 7	2/2
No info	6/20

The following boxes characterize the 31 discordant chemicals (i.e., outliers) for the NHK NRU by counting the number of outliers in each category and comparing to the total number of chemicals in the category.

Boiling Point (in degrees C)	Outliers/Total
No info	13/34
< 100	2/8
100-200	1/5
200-300	3/4
300-400	5/6
408	0/1
960	0/1
1500	0/1
others decompose, sublime, or BPs are given @ < atmospheric pressure	

Molecular Weight (g/mol)	Outliers/Total
< 150	5/21 chemicals (but no info on CCl4)
> 150-200	7/14
200-300	13/20
300-400	3/11
400-500	3/4
500-600	0/1
600-700	0/1
700-800	0/1

Chemical class	Outliers/Total	
Alcohols	4/10	
Carboxylic acids	6/12	
Heterocyclic	9/14	
Mercury compounds	1/1	
Organophosphorous	3/3	(2 were organothiophosphorous cmpds)
Polycyclic	0/2	
Sulfur compound	1/2	
Sodium compound	2/5	
Organometallic	0/1	
Amide	0/3	
Amine	1/1	
Arsenical	1/2	
Boron compounds	0/1	
Cadmium compounds	0/1	
Cyclic hydrocarbon	0/2	
Halogenated hydrocarbon	1/3	
Hydrocarbon	1/2	
Ketone	0/1	
Lithium compound	0/1	
Metal compound	0/1	
Nitrile	0/1	
Phenol	0/1	
Potassium compound	1/2	
Chlorine compound	1/2	
Nitrogen compound	1/1	
Chromium compound	0/1	
Fluorine compound	0/1	
Oxygen compound	1/1	
Selenium compound	1/1	

NHK IC50 (mM)	Outliers/Total
≤ 0.0001	0/2
0.0001 - 0.001	1/2
0.001 - 0.01	3/4
0.01 - 0.1	6/13
0.1 - 1	11/23
1-10	8/12
10-100	1/9
> 100	1/5
No IC50s for 2 chemicals	

pH	Outliers/Total
< 7.1	0/6
7.1	0/0
7.2	1/1
7.3	0/0
7.4	1/4
7.5	3/7
7.6	4/7
7.7	9/23
7.8	11/17
7.9	0/3
8	0/1
8.1	0/0
8.2	1/1
8.3	0/0
8.4	0/0
8.5	1/1
> 8.5	0/1

log Kow	Outliers/Total
< -4	0/1
-4 to –1	2/7
-1 to 0	5/7
0 to 1	3/7
1 to 2	5/13
2 – 3	1/5
3 - 4	5/8
4 - 5	2/2
6-7	1/2
No info	7/20

[This Page Intentionally Left Blank]

Appendix L-2

Discordant Substances for GHS Toxicity Category Predictions Using the 3T3 and NHK NRU Test Methods and Associated Regressions

[This Page Intentionally Left Blank]

L.2 Discordant Substances for GHS Toxicity Category Predictions Using the 3T3 and NHK NRU Test Methods and Associated Regressions

This appendix provides a more detailed discussion of the discordant substances identified for the GHS acute oral toxicity category predictions using the NRU test methods and the regressions evaluated in **Section 6.3.**

L.2.1 Discordant Substances for Prediction of Toxicity Category by the 3T3 and NHK NRU Test Methods and the RC Millimole Regression

Table L2-1 identifies the discordant substances for which the *in vitro* predicted GHS toxicity category (using the 3T3 and NHK NRU test methods with the RC millimole regression) did not match the GHS toxicity category assigned based on the reference rodent LD_{50} data. For the 3T3 NRU test method, the toxicity category was underpredicted for 20 (56%) and overpredicted for 14 (43%) of the 34 discordant substances. Of the 14 substances for which toxicity was underpredicted,

- 8 (57%) were underpredicted by one toxicity category
- 2 (14%) were underpredicted by two toxicity categories
- 4 (29%) were underpredicted by three toxicity categories.

For the 20 substances for which toxicity was overpredicted,

- 14 (70%) were overpredicted by one toxicity category
- 6 (30%) were overpredicted by two toxicity categories.

For the NHK NRU test method, toxicity was underpredicted for 12 (54%) and overpredicted for 22 (46%) of the 34 discordant substances. Of the 12 substances for which toxicity was underpredicted,

- 6 (50%) were underpredicted by one toxicity category
- 3 (25%) were underpredicted by two toxicity categories
- 2 (17%) were underpredicted by three toxicity categories
- 1 (8%) was underpredicted by four toxicity categories

For the 22 substances for which toxicity was overpredicted,

- 16 (73%) were overpredicted by one toxicity category
- 6 (27%) were overpredicted by two toxicity categories

The fact that there were more substances for which toxicity was overpredicted is a result of the removal of substances with specific mechanisms of toxicity that were not expected to be active in the 3T3 and NHK cell cultures. The toxicity for most of these substances would have been underpredicted. **Figure 3-1** shows that most of the 58 selected RC chemicals are below the RC regression line. Thus, the RC would predict lower toxicity (i.e., a higher LD_{50}) for most of these chemicals.

Table L2-1 Discordant Substances[1] for the Prediction of GHS[2] Toxicity Categories by the 3T3 and NHK NRU Test Methods and the RC Millimole Regression

Rodent GHS Toxicity Category[3] (mg/kg)	3T3 NRU Test Method		NHK NRU Test Method	
	Toxicity Overpredicted	Toxicity Underpredicted	Toxicity Overpredicted	Toxicity Underpredicted
$LD_{50} < 5$		Aminopterin (1) Busulfan (3) Cycloheximide (1) Mercury chloride (2) Phenylthiourea (3) Sodium selenate (3) Triethylenemelamine (1)		Aminopterin (4) Busulfan (3) Cycloheximide (1) Mercury chloride (2) Phenylthiourea (3) Sodium selenate (2) Triethylenemelamine (2)
$5 < LD_{50} \leq 50$		Arsenic trioxide (1) Digoxin (3) Sodium arsenite (1) Sodium dichromate dihydrate (1) Thallium sulfate (2)		Arsenic trioxide (1) Sodium dichromate dihydrate (1) Thallium sulfate (1)
$50 < LD_{50} \leq 300$		Cupric sulfate pentahydrate (1) Sodium oxalate (1)	Hexachlorophene (1)	Cupric sulfate pentahydrate (1) Sodium oxalate (1)
$300 < LD_{50} \leq 2000$				
$2000 < LD_{50} \leq 5000$	Acetaminophen (1) Acetonitrile (1) Boric acid (1) Chloramphenicol (1) Citric acid (1) Dimethylformamide (1) Lactic acid (1) Potassium chloride (1) Sodium chloride (1) Trichloroacetic acid (1) Xylene (1)		Acetaminophen (1) Acetonitrile (1) Boric acid (1) Chloramphenicol (1) Citric acid (1) Lactic acid (1) Potassium chloride (1) Sodium chloride(1) Trichloroacetic acid (1) Xylene (1)	
$LD_{50} > 5000$	5-Aminosalicylic acid (2) Dibutyl phthalate (2) Diethyl phthalate (2) Ethanol (2) Ethylene glycol (1)		5-Aminosalicylic acid (2) Dibutyl phthalate (2) Diethyl phthalate (2) Ethanol (1) Ethylene glycol (1)	

Rodent GHS Toxicity Category[3] (mg/kg)	3T3 NRU Test Method		NHK NRU Test Method	
	Toxicity Overpredicted	**Toxicity Underpredicted**	**Toxicity Overpredicted**	**Toxicity Underpredicted**
	Glycerol (1) 2-Propanol (2) Sodium hypochlorite (2) 1,1,1-Trichloroethane (1)		Gibberellic acid (1) Glycerol (1) Methanol (2 2-Propanol (2) Sodium hypochlorite (2) 1,1,1-Trichloroethane (1)	

[1]Substances for which the *in vitro* predicted GHS toxicity category was different from the GHS toxicity category assigned to the substance based on reference rodent LD_{50} data. Numbers in parentheses indicate the number of categories different. Three substances were excluded because no rat LD_{50} was identified: epinephrine bitartrate, colchicine, and propylparaben. Carbon tetrachloride was excluded from the 3T3 and NHK NRU analyses because no laboratory attained sufficient toxicity for the calculation of an IC_{50}. Methanol was excluded from the 3T3 analysis because no laboratory attained sufficient toxicity for the calculation of an IC_{50}. The 21 substances in **Table 6-3** were excluded based on their mechanisms of action.
[2]GHS-Globally Harmonized System of Classification and Labelling of Chemicals with LD_{50} in mg/kg (UN 2005). The RC millimole regression is log LD_{50} (mmol/kg) = log IC_{50} (mM) X 0.435 + 0.625.
[3]Reference rodent LD_{50} values in from **Table 3-2**.

L.2.2 Discordant Substances for Prediction of Toxicity Category by the 3T3 and NHK NRU Test Methods and the RC Rat-Only Weight Regression

Table L2-2 shows the discordant substances for which the *in vitro* predicted GHS toxicity category (using the 3T3 and NHK NRU test methods with the RC rat-only weight regression) did not match that based on the reference rodent LD_{50} data. The two *in vitro* NRU cytotoxicity test methods over- and under-predicted the GHS toxicity category for a similar number of substances. For the 3T3 NRU test method, the GHS toxicity category of 19 (63%) of 30 discordant substances was overpredicted, with:

- 13 (68%) overpredicted by one GHS toxicity category
- 6 (32%) overpredicted by two GHS toxicity categories

The toxicity of fewer substances (11; 37%) was underpredicted by this test method, with:

- 7 (64%) underpredicted by one GHS toxicity category
- 4 (36%) underpredicted by two GHS toxicity categories

For the NHK NRU test method, the GHS toxicity category of 22 (67%) of the 33 discordant substances was overpredicted. Of these,

- 15 (68%) were overpredicted by one GHS toxicity category
- 7 (31%) were overpredicted by two GHS toxicity categories

For this assay, the toxicity of 11 (33%) of the discordant substances was underpredicted, with

- 6 (55%) underpredicted by one GHS toxicity category
- 4 (36%) underpredicted by two GHS toxicity categories
- 1 (9%) underpredicted by three toxicity categories.

Phenylthiourea, the substance for which the GHS toxicity category was underpredicted by three toxicity categories, was in the most severe GHS toxicity category (i.e., $LD_{50} < 5$ mg/kg).

Table L2-2 Discordant substances[1] for RC Rat-Only Weight Regression Prediction of GHS Toxicity Categories[2] by the 3T3 and NHK NRU Test Methods

Rodent GHS Toxicity Category[3] (mg/kg)	3T3 NRU Test Method		NHK NRU Test Method	
	Toxicity Overpredicted	Toxicity Underpredicted	Toxicity Overpredicted	Toxicity Underpredicted
$LD_{50} < 5$		Cycloheximide (1) Phenylthiourea (2) Sodium selenate (2) Triethylenemelamine (1)		Cycloheximide (1) Phenylthiourea (3) Sodium selenate (2) Triethylenemelamine (2)
$5 < LD_{50} \leq 50$		Arsenic trioxide (1) Busulfan (2) Digoxin (2) Mercury chloride (1) Thallium sulfate (1) Sodium arsenite (1)		Aminopterin (2) Arsenic trioxide (1) Busulfan (2) Mercury chloride (1) Sodium arsenite (1) Thallium Sulfate (1)
$50 < LD_{50} \leq 300$		Sodium fluoride (1)	Hexachlorophene (1)	Sodium fluoride (1)
$300 < LD_{50} \leq 2000$	Propranolol (1) Triphenyltin hydroxide (2)		Triphenyltin hydroxide (2)	
$2000 < LD_{50} \leq 5000$	Acetaminophen (1) 5-Aminosalicylic acid (1) Boric acid (1) Chloramphenicol (1) Xylene (1)		5-Aminosalicylic acid (1) Chloramphenicol (1) Boric acid (1) Xylene (1) Acetaminophen (1) Lactic acid (1) Sodium chloride (1) Potassium chloride (1)	
$LD_{50} > 5000$	Citric acid (2) Diethyl phthalate (2) Dibutyl phthalate (2) Dimethylformamide (1) Ethanol (1) Ethylene glycol (1) Gibberellic acid (1) Glycerol (1) 2-Propanol (1)		Citric acid (2) Dibutyl phthalate (2) Diethyl phthalate (2) Dimethylformamide (1) Ethanol (1) Gibberellic Acid (1) Glycerol (1) Methanol (2) 2-Propanol (1)	

Rodent GHS Toxicity Category[3] (mg/kg)	3T3 NRU Test Method		NHK NRU Test Method	
	Toxicity Overpredicted	**Toxicity Underpredicted**	**Toxicity Overpredicted**	**Toxicity Underpredicted**
	Sodium hypochlorite (2) Trichloroacetic acid (2) 1,1,1-Trichloroethane (1)		Sodium hypochlorite (2) Trichloroacetic acid (2) 1,1,1-Trichloroethane (1)	

[1]Substances for which the *in vitro* predicted GHS toxicity category was different from that based on the reference rodent LD_{50} data. Numbers in parentheses indicate the number of categories different. Three substances were excluded because no rat LD_{50} was identified: epinephrine bitartrate, colchicine, and propylparaben. Carbon tetrachloride was excluded from the 3T3 and NHK NRU analyses because no laboratory attained sufficient toxicity for the calculation of an IC_{50}. Methanol was excluded from the 3T3 NRU analysis because no laboratory attained sufficient toxicity for the calculation of an IC_{50}. The 21 substances in **Table 6-3** were excluded based on their mechanisms of action.

[2]GHS-Globally Harmonized System of Classification and Labelling of Chemicals with LD_{50} in mg/kg (UN 2005). The RC rat-only weight regression is log LD_{50} (mg/kg) = log IC_{50} (µg/mL) X 0.372 + 2.024.

[3]Reference rodent LD_{50} values from **Table 4-2**.

L.2.3 Discordant Substances for the Prediction of Toxicity Category by the 3T3 and NHK NRU Test Methods and the RC Rat-Only Weight Regression Excluding Substances with Specific Mechanisms of Toxicity

Table L2-3 shows the discordant substances for which the *in vitro* NRU predicted toxicity category (using the 3T3 and NHK NRU test methods with the RC rat-only weight regression excluding substances with specific mechanisms of toxicity) did not match that based on the reference rodent LD_{50} data. The NHK NRU test method had four more discordant substances than the corresponding assay using 3T3 cells when the IC_{50} results were applied to the RC rat-only weight regression, after excluding substances with specific mechanisms of toxicity. For the 3T3 NRU test method, the GHS toxicity category of 19 (63%) of 30 discordant substances was overpredicted, with

- 13 (68%) overpredicted by one toxicity category
- 6 (32%) overpredicted by two toxicity categories

The toxicity of 11 (37%) of 30 discordant substances was underpredicted by the 3T3 NRU test method, with

- 7 (64%) underpredicted by one category
- 4 (36%) underpredicted by two toxicity categories

For the NHK NRU test method, the toxicity of 22 (65%) of 34 discordant substances was overpredicted, with

- 15 (68%) overpredicted by one category
- 7 (32%) overpredicted by two toxicity categories

Also, for this test method, the toxicity of 12 (35%) of 34 discordant substances was underpredicted, with

- 6 (50%) underpredicted by one toxicity category
- 4 (33%) underpredicted by two toxicity categories
- 2 (17%) underpredicted by three toxicity categories

Both substances for which toxicity was underpredicted by three toxicity categories were in the most severe GHS toxicity category (i.e., LD_{50} < 5 mg/kg).

Table L2-3 Discordant Substances[1] for RC Rat-Only Weight Regression Excluding Chemicals with Specific Mechanisms of Toxicity Prediction of GHS Toxicity Categories[2] by the 3T3 and NHK NRU Test Methods

Rodent GHS Toxicity Category[3] (mg/kg)	3T3 NRU Test Method		NHK NRU Test Method	
	Toxicity Overpredicted	Toxicity Underpredicted	Toxicity Overpredicted	Toxicity Underpredicted
$LD_{50} < 5$		Cycloheximide (1) Phenylthiourea (2) Sodium selenate (2) Triethylenemelamine (1)		Cycloheximide (1) Disulfoton (3) Phenylthiourea (3) Sodium selenate (2) Triethylenemelamine (2)
$5 < LD_{50} \leq 50$		Arsenic trioxide (1) Busulfan (2) Digoxin (2) Mercury chloride (1) Thallium sulfate (1) Sodium arsenite (1)		Aminopterin (2) Arsenic trioxide (1) Busulfan (2) Mercury chloride (1) Sodium arsenite (1) Thallium sulfate (1)
$50 < LD_{50} \leq 300$		Sodium fluoride (1)	Hexachlorophene (1)	Sodium fluoride (1)
$300 < LD_{50} \leq 2000$	Propranolol (1) Triphenyltin hydroxide (2)		Triphenyltin hydroxide (2)	
$2000 < LD_{50} \leq 5000$	Acetaminophen (1) 5-Aminosalicylic acid (1) Boric acid (1) Chloramphenicol (1) Xylene (1)		5-Aminosalicylic acid (1) Chloramphenicol (1) Boric acid (1) Xylene (1) Acetaminophen (1) Lactic acid (1) Sodium chloride (1) Potassium chloride (1)	
$LD_{50} > 5000$	Citric acid (2) Diethyl phthalate (2) Dibutyl phthalate (2) Dimethylformamide (1) Ethanol (1) Ethylene glycol (1) Gibberellic acid (1)		Citric acid (2) Dibutyl phthalate (2) Diethyl phthalate (2) Dimethylformamide (1) Ethanol (1) Gibberellic acid (1) Glycerol (1)	

Rodent GHS Toxicity Category[3] (mg/kg)	3T3 NRU Test Method		NHK NRU Test Method	
	Toxicity Overpredicted	**Toxicity Underpredicted**	**Toxicity Overpredicted**	**Toxicity Underpredicted**
	Glycerol (1) 2-Propanol (1) Sodium hypochlorite (2) Trichloroacetic acid (2) 1,1,1-Trichloroethane (1)		Methanol (2) 2-Propanol (1) Sodium hypochlorite (2) Trichloroacetic acid (2) 1,1,1-Trichloroethane (1)	

[1]Substances for which the *in vitro* predicted GHS toxicity category was different from that based on the reference rodent LD_{50} data. Numbers in parentheses indicate the number of categories different. Three substances were excluded because no rat LD_{50} was identified: epinephrine bitartrate, colchicine, and propylparaben. Carbon tetrachloride was excluded from the 3T3 and NHK NRU analyses because no laboratory attained sufficient toxicity for the calculation of an IC_{50}. Methanol was excluded from the 3T3 NRU analysis because no laboratory attained sufficient toxicity for the calculation of an IC_{50}. The 21 substances in **Table 6-3** were excluded based on their mechanisms of action.

[2]GHS-Globally Harmonized System of Classification and Labelling of Chemicals with LD_{50} in mg/kg (UN 2005). The RC rat minus specific mechanisms regression is log LD_{50} (mg/kg) = log IC_{50} (μg/mL) X 0.357 + 2.194.

[3]Reference rodent LD_{50} values from **Table 4-2**.

[This Page Intentionally Left Blank]

Appendix L-3

Analysis of Outliers by Halle (1998) for the RC Millimole Regression

[This Page Intentionally Left Blank]

L.3 Analysis of Outliers for the RC Millimole Regression

The RC millimole regression was constructed from the *in vitro* IC_{50X} cytotoxicity data from multiple cell lines and the *in vivo* acute toxicity data from rats and mice (i.e., LD_{50} values) for 347 chemicals (Halle 1998). Halle (1998) investigated the 95 (27.4%) chemicals for which the observed log LD_{50} values were greater than 0.699 (i.e., 0.5 log) from predicted log LD_{50} values. Of the 95 outliers, 46 were positive outliers and 49 were negative outliers. The positive outliers have IC_{50X} values that predict a far higher *in vivo* toxicity (i.e., lower LD_{50}) than the actual animal experiment. The negative outliers are more important since the IC_{50X} values predict lower toxicity (i.e. higher LD_{50}) than the observed *in vivo* toxicity. It seems that Halle (1998) was not concerned about the positive outliers since the prediction erred in a health protective direction. Halle (1998) was much more concerned about trying to explain the reasons for the negative outliers since the error was in a nonconservative direction.

Halle (1998) investigated three factors that could have explained the negative outliers.

1. Variation in the oral LD_{50} values

   Reported oral LD_{50} values for a particular chemical might vary by factor 4 to 14 even when experiments were highly standardized. They found LD_{50} values from other sources for 23 of the 95 outliers. They found that the variations in the LD_{50} values (difference between the RTECS® value and the "new" value found for the 23 chemicals) were larger for the negative outliers than for the positive outliers.

2. Species-specificity of the oral LD_{50} values.

   Halle (1998) compared an IC_{50x}–LD_{50} regression using mouse LD_{50} values (242 values) with a regression using rat LD_{50} values (285 values) and found no significant difference between the two regressions. The RC millimole regression with 347 chemicals has 285 rat values and 65 mouse values and is not statistically different from either the rat or mouse regressions.

3. The cell culture(s) used may have been unsuitable for the detection of cytotoxic potential or it may have been unable to simulate the complex process of toxicity *in vivo.*

Halle (1998) expected, *a priori*, that three classes of compounds, insecticides (**Table L3-1**), neurotoxins (**Table L3-2**), and those requiring metabolic activation for toxicity **(Table L3-3**), would not fit the RC millimole regression (i.e., cytotoxicity data would not predict *in vivo* toxicity). Sixty-two of the 347 chemicals belong to these three classes. Twenty-three (37.1%) of the 62 chemicals were negative outliers. Of the 23, 10 were insecticides, 5 were neurotoxins, and 8 required metabolic activation. No positive outliers were identified in the three classes.

Of the 49 negative outliers, 23 (46.9%) belonged to the three classes of concern. Examination of these classes showed that the RC millimole prediction was accurate (i.e., predicted log LD_{50} [mmol/kg] was within 0.699 of observed log LD_{50} in [mmol/kg]) for 50% of the insecticides (**Table L3-1**) and chemicals that required metabolic activation (**Table L3-3**). For neurotoxins (**Table L3-2**), the results were even better, since 21 (80.8%) fell within the prediction interval. Halle (1998) felt that the ability to predict the acute LD_{50} for 50% of the insecticides and xenobiotics requiring metabolic activation and for 81% of the neurotoxic xenobiotics was sufficiently accurate for practical purposes.

Of the 49 negative outliers in the RC millimole regression, 23 (46.9%) of these belonged to the three classes of concern that may explain the false negative IC_{50X} values. Findings were contrary to Halle's assumption that *in vitro* cytotoxicity would not predict *in vivo* toxicity for these types of chemicals. The RC millimole prediction of LD_{50} was applicable to 50% of the insecticides and chemicals that required metabolic activation. For neurotoxic chemicals the results were even better, since 21 (80.8%) fell within the prediction interval. Halle felt that the ability to predict the acute LD_{50} for 50% of the insecticides and chemicals requiring metabolic activation and for 81% of the neurotoxic chemicals was sufficiently accurate for practical purposes.

In separate analyses, Halle (1998) considered the physicochemical properties of chemicals (i.e., molecular weight and the octanol/water partition coefficient) as independent variables in a multiple regression analysis, but they did not improve the prediction of LD_{50} by IC_{50}.

L3-1 The Error of Prediction[a] of 20 of The Most Important Insecticides in the RC Ordered According to Their Chemical Characteristics[b]

Chemical class	RC No	Name	LD_{50} Error of Prediction[a]
Chlorinated hydrocarbon			
	26	Kelthane	0.340
	40	Chlordan	-0.046
	43	Aldrin	**-1.074**[b]
	61	DDT	**-0.775**
	167	DDD	-0.378
	185	Heptachlor	**-1.050**
	195	DDA	0.133
	197	DDE	0.251
	207	Dieldrin	**-1.223**
	223	Lindane	**-1.043**
Organophosphorus compounds			
	49	Parathion	**-2.339**
	51	Disulfoton	**-2.346**
	67	Malathion	0.106
	75	Trichlorfon	-0.136
	96	Cygon	**-0.848**
Carbamate compounds			
	73	Carbaryl	-0.279
	186	Zineb	**1.185**
Other compounds			
	134	Rotenone	0.583
	173	Pentachlorophenol	**-0.720**
	235	Paraquat	**-1.019**

[a] defined as observed log LD_{50} (mmol/kg) - predicted log LD_{50} (mmol/kg)
[b] modified from Table 10 of Halle (1998)
bold numbers: outliers (i.e., observed log LD_{50} (mmol/kg) - predicted log LD_{50} (mmol/kg) > 0.699)

Table L3-2 The Error of Prediction [a] of 26 Neurotoxic Xenobiotics in the RC Ordered According to Their *In Vivo* Potency[b]

Chemical Class	RC No	Name	LD_{50} Error of Prediction[a]
Sedative, hypnotic, CNS depressants			
	69	Secobarbital sod.	-0.651
	83	Thiopental	-0.119
	84	Amobarbital	-0.335
	87	Pentobarbital sodium	-0.654
	101	Gluthetimide	-0.270
	118	Phenobarbital	**-1.035**[b]
	247	(+)-Thalidomide	-0.397
	264	Chloral hydrate	-0.349
	317	Barbital sodium	-0.591
Antidepressant			
	38	Imipramine HCl	-0.093
	90	Iproniazid	-0.273
	183	Amitriptyline	0.021
Antipsychotic, anxiolytic			
	27	Chlorpromazine	-0.176
	44	Hydroxyzine HCl	0.248
	63	Diazepam	0.116
	170	Thioridazine HCl	-0.013
Stimulants			
	112	Caffeine	**-0.815**
	262	Amphetamine sulfate	**-1.579**
Anticonvulsants			
	82	Diphenylhydantoin	-0.551
Analgetic (general anesthesia)			
	229	Dextropropoxyphene HCl	**-1.150**
Anticholinergic			
	251	Scopolamine * HBr	-0.123
	296	Homatropine methylbromide	-0.532
Other Neurotoxins (not insecticide)			
	102	Acrylamide	-0.338
	137	Triethyltin chloride	**-0.852**
	142	Methylmercury chloride	0.105
	316	Toluene	0.571

[a] defined as observed log LD_{50} (mmol/kg) - predicted log LD_{50} (mmol/kg)
[b] modified from Table 11 of Halle (1998)
bold numbers: outliers (i.e., observed log LD_{50} (mmol/kg) - predicted log LD_{50} (mmol/kg) > 0.699)

Table L3-3 The Error of Prediction [a] of the 16 Xenobiotics in the RC that Require Metabolic Activation [b]

RC No	Name	LD_{50} error of prediction[a]
13	Cycloheximide	**-1.370**[b]
33	p-Chloromercuribenzoic acid	**-1.077**
37	Aflatoxin B_1	**-1.783**
68	2.4-Dinitrophenol	**-1.128**
97	Phenacetin	0.292
109	Frusemide	0.109
113	Acetaminophen	0.386
116	Cyclophosphamide * H_2O	**-1.310**
123	Isoniazid	-0.332
125	Carbon tetrachloride	0.229
192	1.3-Bis(2-chloroethyl)-1-nitrosourea	**-1.176**
260	Coumarin	-0.427
273	Bromobenzene	0.374
279	Thioacetamide	-0.294
281	1.2-Dibromomethane	**-1.106**
292	Allylalcohol	**-0.952**

[a] defined as observed log LD_{50} (mmol/kg) - predicted log LD_{50} (mmol/kg)

[b] modified from Table 12 of Halle (1998)

bold numbers: outliers (i.e., observed log LD_{50} (mmol/kg) - predicted log LD_{50} (mmol/kg) > 0.699

[This Page Intentionally Left Blank]