III. Appendices

B. Hazard/RPF

3. Response to SAP Comments from September 2001 and March 2002 Reports

a. Response to SAP Comments from September 2001

OPP in collaboration with ORD presented its July 31st , 2001 document entitled, "Determination of Relative Potency and Points of Departure for Cholinesterase Inhibition" to the FIFRA SAP on September 5-6, 2001. The key recommendations from the September 2001 report (http://www.epa.gov/scipoly/sap/index.htm) and OPP's responses are given below:

i. Derivation of the Adjustment Factor "B" and Modification of Decision Tree for use of "B"

The SAP Report noted that a plot of the "scaled residuals" against "predicted % inhibition" indicates that the weighting strategy used for calculating the adjustment factor "B" does not adequately reflect how the variance changes with response. The SAP was specifically concerned EPA "focused the modeling effort on achieving fidelity with observations at the high end of the range of doses tested, to the likely detriment of fitting points at the low end of the dose response relationship."

In the current analysis, all available cholinesterase datasets for the brain compartment were analyzed using a fixed horizontal y-asymptote for each chemical. The weight function was changed from one in which the variance was presumed proportional to the square of the mean to one in which the variance is proportional to the mean. The revised methodology for the determination of the horizontal y- asymptote is described in I.B and III.B.1.

ii. Conduct a Formal Analysis of Residuals as a Function of Dose

Residual plots for the basic and expanded models for each chemical for the brain compartment are given in III.B.2.

iii. Accuracy of the "Chi Square Approximation" for the "Goodness of Fit" Statistic

In the July 31st document, a Chi-Square Approximation was calculated for each cholinesterase dataset. This statistic was used as a measure of the goodness-of-fit for the exponential function. The concern expressed by the SAP does not apply to the current methodology. Although the OPCumRisk program was not used to determine potency of OPs in the current analysis, the program was revised to deliver a warning message to the program user indicating possible calculation inaccuracy for this statistic. The revised version of the OPCumRisk is available for download at http://www.epa.gov/scipoly/sap/index.htm and http://www.epa.gov/pesticides/cumulative/

iv. Confidence Interval Calculations

The SAP report suggested that HED "reconsider the confidence interval calculations"and "perhaps try bootstrapping or some other more robust method . . . ." In the current analysis, HED has revised the calculation of the confidence intervals (See III.B.1). Bootstrapping is a very time and resource-intensive procedure. Although bootstrapping may be the preferred approach for calculating confidence intervals, due to limited availability of resources, the Agency has not conducted any bootstrapping procedures. At this time, the current method for calculating confidence intervals is adequate and satisfactory. Because it is important to evaluate the range of uncertainty around any potency or benchmark dose values used to extrapolate to human risk, the Agency will consider bootstrapping procedures in future assessments.

v. Deleting p- and t- values

The SAP Report recommended deleting the p- and t- values that are produced by the Agency's OPCumRisk program. As stated previously, the OPCumRisk program was not used in the current analysis to calculate potency or benchmark dose estimates. The requested deletions have been incorporated; the revised version of the OPCumRisk is available for download at http://www.epa.gov/scipoly/sap/index.htm and http://www.epa.gov/pesticides/cumulative/

vi. Estimates of Relative Potency

The SAP Report included considerable discussion regarding whether relative potency factors should be based on ratios of the “Benchmark Dose 10's” (BMD_{10}) or on ratios of the dose-scaling factors. OPP has derived potency in the present analysis on BMD_{10} (See I.B).

vii. Inhalation Dose

The SAP Report recommended that inhalation exposure be expressed in the same units as the oral doses and that the doses be adjusted for actual treatment durations. HED has calculated the inhalation doses as mg/kg/day using conversion factors that account for respiratory volume and body weight for the strain of rat used, as well as the duration of exposure in terms of hours exposed per day.

viii. Use of Individual Animal Data

The SAP Report from the September 2000 SAP meeting recommended that study data on individual animals be used in calculating relative potencies. Due to the fact that all the data on organophosphates are not in an electronic format, HED has not taken this step. However, the September, 2001 Report recognizes that "individual data would not be likely to change the results using current methods." In addition, by switching from RBC to the brain compartment, some of the concern about not using individual animal data should be reduced, since the experimental designs for the brain measurements do not include a repeated measures component, unlike the RBC data.

iv. Use of NOAEL's and LOAEL's for Inhalation and Dermal Routes

Several Panel members objected to EPA's use of No Observed Adverse Effect Levels ("NOAEL's") and Lowest Observed Adverse Effect Levels ("LOAEL's") for cholinesterase inhibition data by the dermal and inhalation routes of exposure instead of actual dose-response models as are used for the oral data set. HED does not intend to use dose-response modeling to determine relative potency estimates for dermal and inhalation exposure because the data are not sufficiently robust to justify the resources required.

However, it is to be noted that the current analysis uses Comparative Effect Levels (CEL's) for cholinesterase inhibition data for these two routes of exposure. The dermal and inhalation database was not suitable for dose-response analysis. Cholinesterase determinations in these studies were typically made at only one time point and several of the studies had no cholinesterase inhibition at the highest dose. For the current assessment, potencies by the dermal and inhalation routes were compared using brain cholinesterase inhibition at a dose causing a maximum of 15% brain cholinesterase inhibition.

v. Derivation of Doses from the Actual Dietary Intake Rates

The SAP Report recommends that "the doses used for evaluation of potencies at various ages within specific data sets should be derived from the actual dietary intake rates observed in the study for those ages where the consumption data are available."

In feeding toxicity studies, laboratory rats are exposed to the test compound via the diet. Generally, the test compound is mixed in the animal feed which the laboratory animals eat. Over the course of a toxicity study, as the animals age, they will not only gain weight and but they will naturally change their rate of food consumption. The data collected for the oral route and used in both the July and December 2001 preliminary cumulative risk assessments include average compound intake (mg of active ingredient per kg per day). HED has conducted a pilot analysis in response to this recommendation to evaluate the effect of age and food consumption rate on the potency estimates. In this pilot compound intake analysis, OP potency was determined for a subset of studies [≈10% of total studies in the dose-response assessment] using compound intake measured at or around the time of cholinesterase measurements [duration-specific compound intake].

Seventy-nine oral toxicity studies were included in the dose-response assessment for the December, 2001 Cumulative Risk Assessment for OPs. Of these 79 studies, the test article was administered via the diet for 73. For each of the seven OPs selected for this analysis, the calculated compound intake (mg/kg/day) given in the study report for a weekly, biweekly, or monthly time interval closest to the time of cholinesterase measurement was extracted from the feeding toxicity studies [duration-specific compound intakes]. For example, if brain cholinesterase was measured at a one-year interim sacrifice, the compound intake for the 50-52 week reported interval was collected. The potency values obtained were compared to those in the July, 2001 analysis, which utilized average compound intake values. Potency estimates given below (Table III.B.3-4) were calculated using the OPCumRisk program with the methodology described in the July 31st document *prior* to the completion of the current methodology for the joint analysis. The pilot analysis was performed in three stages : 1) impact of age on relative potency for chronic studies only; 2) impact of age on relative potency for complete database of subchronic and chronic studies; and 3) impact of age on the points of departure on the index chemical.

Stage 1: The purpose of this pilot analysis was to investigate the impact of age on food consumption and body weight, and ultimately OP potency. In order to maximize the age-related differences in body weight and food consumption, chronic studies were analyzed first. Seven chronic feeding studies were selected

randomly and analyzed as described above. Relative potency of each was calculated using the methamidophos chronic study. Results given in Table III.B.3-1.

In the chronic study analysis (Table III.B.3-1) comparing the RPFs calculated using the slope scale factor (m) and also the BMD_{10}s for ChE data using the average and duration-specific compound intakes, *the RBC and brain data for both sexes display comparable potency values*. For tribufos a 5-fold difference between the average and duration-specific intake assessments for male brain CHeI was observed. This difference is an artifact of the decision tree for the determination B (horizontal asymptote) and not from differences in potency between the average and duration specific intakes. Two timepoints (364 and 721 days) are available for the male brain ChE data in MRID 42335101. In the duration specific analysis, the 364 day time point did not converge and was therefore not included in the potency estimates.

Table III.B.3-1a. Results of Dietary Intake Comparison [actual *vs* average] Using Chronic Studies

CHEMICAL	MRID	COMPARTMENT	SEX	Dietary Intake Calculation	Relative Potency using 'm'	Lower 95% CL	Upper 95% CL	BMD_{10}	BMDL	Relative Potency using BMD_{10}
BENSULIDE	44161101	BRAIN	F	average	0.005	0.004	0.006	14.11	12.40	0.005
				biweekly	0.004	0.004	0.005	14.04	12.17	0.004
DIAZINON	41942002	BRAIN	F	average	0.034	0.031	0.038	1.85	1.78	0.038
				biweekly	0.031	0.028	0.035	1.85	1.80	0.034
DICROTOPHOS	44527802	BRAIN	F	average	1.77	1.41	2.22	0.041	0.035	1.74
				biweekly	1.89	1.51	2.38	0.035	0.030	1.79
METHAMIDOPHOS	00148452	BRAIN	F	average	1.00	1.00	1.00	0.071	0.063	1.00
				biweekly	1.00	1.00	1.00	0.063	0.058	1.00
PHOSALONE	44801002	BRAIN	F	average	0.015	0.013	0.018	4.13	3.70	0.017
				biweekly	0.024	0.020	0.029	2.40	2.14	0.026
PHOSMET	41916401	BRAIN	F	average	0.023	0.010	0.053	4.41	3.74	0.016
				biweekly	0.021	0.016	0.027	2.76	2.33	0.023
TRIBUFOS	42335101	BRAIN	F	average	0.018	0.007	0.048	3.26	1.88	0.022
				biweekly	0.017	0.007	0.045	3.14	1.83	0.020
BENSULIDE	44161101	BRAIN	M	average	0.002	0.002	0.003	24.69	19.37	0.003
				biweekly	0.002	0.001	0.003	24.93	19.54	0.002
DIAZINON	41942002	BRAIN	M	average	0.011	0.003	0.041	3.38	1.83	0.018
				biweekly	0.011	0.003	0.035	3.31	1.83	0.016
DICROTOPHOS	44527802	BRAIN	M	average	2.06	1.70	2.38	0.028	0.026	2.23
				biweekly	2.32	2.03	2.67	0.022	0.020	2.45
METHAMIDOPHOS	00148452	BRAIN	M	average	1.00	1.00	1.00	0.062	0.057	1.00
				biweekly	1.00	1.00	1.00	0.055	0.049	1.00
PHOSALONE	44801002	BRAIN	M	average	0.021	0.018	0.025	2.58	2.37	0.024
				biweekly	0.038	0.033	0.044	1.29	1.18	0.042
PHOSMET	41916401	BRAIN	M	average	0.011	0.008	0.015	5.35	4.33	0.012
				biweekly	0.013	0.009	0.018	3.71	2.98	0.015
TRIBUFOS	42335101	BRAIN	M	average	0.020	0.017	0.022	4.22	2.51	0.015
				biweekly	0.004	0.001	0.020	15.64	6.19	0.003

Table III.B.3-1b. Results of Dietary Intake Comparison [actual *vs* average] Using Chronic Studies

CHEMICAL	MRID	COMPARTMENT	SEX	Dietary Intake Calculation	Relative Potency using 'm'	Lower 95% CL	Upper 95% CL	BMD_{10}	BMDL	Relative Potency using BMD_{10}
BENSULIDE	44161101	RBC	F	average	0.012	0.005	0.025	5.53	3.69	0.012
				biweekly	0.011	0.005	0.024	5.35	3.55	0.012
DIAZINON	41942002	RBC	F	average	0.12	0.037	0.38	0.28	0.17	0.24
				biweekly	0.11	0.036	0.33	0.29	0.18	0.21
DICROTOPHOS	44527802	RBC	F	average	2.77	1.88	4.08	0.039	0.030	1.71
				biweekly	2.89	1.95	4.29	0.035	0.027	1.78
METHAMIDOPHOS	00148452	RBC	F	average	1.00	1.00	1.00	0.067	0.063	1.00
				biweekly	1.00	1.00	1.00	0.062	0.058	1.00
PHOSALONE	44801002	RBC	F	average	0.068	0.027	0.17	0.71	0.48	0.094
				biweekly	0.076	0.035	0.17	0.64	0.44	0.097
PHOSMET	41916401	RBC	F	average	0.080	0.058	0.11	0.84	0.75	0.080
				biweekly	0.083	0.065	0.11	0.70	0.57	0.089
TRIBUFOS	42335101	RBC	F	average	0.095	0.048	0.19	0.61	0.48	0.11
				biweekly	0.089	0.045	0.18	0.60	0.46	0.10
BENSULIDE	44161101	RBC	M	average	0.013	0.006	0.026	7.56	6.34	0.008
				biweekly	0.013	0.006	0.027	7.55	6.33	0.007
DIAZINON	41942002	RBC	M	average	0.040	0.013	0.13	2.36	1.92	0.025
				biweekly	0.042	0.013	0.13	2.09	1.57	0.025
DICROTOPHOS	44527802	RBC	M	average	1.33	1.10	1.61	0.039	0.035	1.51
				biweekly	1.55	1.26	1.91	0.033	0.030	1.60
METHAMIDOPHOS	00148452	RBC	M	average	1.00	1.00	1.00	0.059	0.056	1.00
				biweekly	1.00	1.00	1.00	0.053	0.047	1.00
PHOSALONE	44801002	RBC	M	average	0.053	0.021	0.13	0.96	0.56	0.062
				biweekly	0.067	0.032	0.14	1.49	1.31	0.035
PHOSMET	41916401	RBC	M	average	0.079	0.055	0.11	0.81	0.72	0.073
				biweekly	0.10	0.077	0.14	0.58	0.53	0.091
TRIBUFOS	42335101	RBC	M	average	0.14	0.090	0.21	0.49	0.40	0.12
				biweekly	0.10	0.050	0.21	0.57	0.42	0.094

Stage 2: Out of the seven OPs analyzed in Stage 1, the entire oral databases; i.e., both chronic and subchronic studies, of three randomly selected OPs were analyzed as in Stage 1. Relative potency was calculated using all available methamidophos studies (Table III.B.3-2).

In the pilot analysis of the complete oral database for three OPs (diazinon, dimethoate, and phosalone; Table III.B.3-2) comparing the RPFs calculated with slope scale factors and BMD_{10}s for ChE data using the average and duration-specific compound intakes, *the RBC and brain data for both sexes display comparable potency values*. For phosalone RBC male *only*, a 7-fold difference between the average and duration-specific intake assessments was observed.

Graphs of potency vs. time are shown in Figures III.B.3-1,2 for the analyzes of average chemical intake and for duration specific chemical intake. The patterns observed in the graphs for the average intake analyzes are similar to those of the duration specific intakes.

Table III.B.3-2. Results of Dietary Intake [actual *vs* average] Using All Available Studies

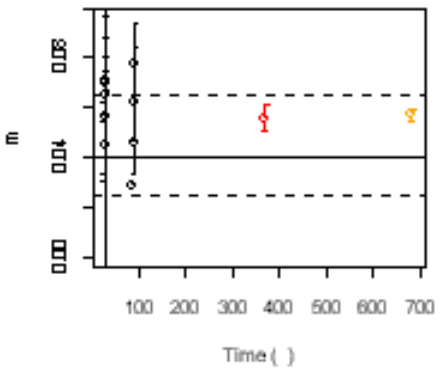
CHEMICAL	MRID	COMPARTMENT	SEX	Dietary Intake Calculation	Relative Potency using 'm'	Lower 95% CL	Upper 95% CL	BMD_{10}	BMDL	Relative Potency using BMD_{10}
DIAZINON	43543901 43543902 40815003 41942002	BRAIN	F	average	0.031	0.018	0.053	2.48	1.78	0.036
				biweekly	0.033	0.019	0.058	2.08	1.51	0.038
DIMETHOATE	43128201 164177	BRAIN	F	average	0.531	0.41	0.69	0.25	0.23	0.36
				biweekly	0.58	0.45	0.75	0.20	0.18	0.40
METHAMIDOPHOS	41867201 00148452 43197901	BRAIN	F	average	1.00	1.00	1.00	0.09	0.08	1.00
				biweekly	1.00	1.00	1.00	0.08	0.07	1.00
PHOSALONE	44852504 44801002	BRAIN	F	average	0.019	0.014	0.025	5.05	3.83	0.018
				biweekly	0.021	0.010	0.040	3.37	2.24	0.024
DIAZINON	43543901 43543902 40815003 41942002	BRAIN	M	average	0.005	0.002	0.012	24.77	24.15	0.003
				biweekly	0.005	0.002	0.010	18.28	17.83	0.004
DIMETHOATE	43128201 164177	BRAIN	M	average	0.71	0.53	0.94	0.10	0.08	0.80
				biweekly	0.83	0.60	1.15	0.08	0.06	0.88
METHAMIDOPHOS	41867201 148452 43197901	BRAIN	M	average	1.00	1.00	1.00	0.08	0.07	1.00
				biweekly	1.00	1.00	1.00	0.07	0.06	1.00
PHOSALONE	44852504 44801002	BRAIN	M	average	0.019	0.011	0.032	3.49	2.49	0.023
				biweekly	0.028	0.012	0.063	1.96	1.22	0.036
DIAZINON	43543901 43543902 40815003 41942002	RBC	F	average	0.38	0.22	0.65	0.24	0.22	0.38
				biweekly	0.41	0.27	0.62	0.18	0.17	0.44
DIMETHOATE	43128201 164177	RBC	F	average	0.32	0.14	0.73	0.29	0.14	0.31
				biweekly	0.27	0.14	0.53	0.33	0.16	0.24

CHEMICAL	MRID	COMPARTMENT	SEX	Dietary Intake Calculation	Relative Potency using 'm'	Lower 95% CL	Upper 95% CL	BMD_{10}	BMDL	Relative Potency using BMD_{10}
METHAMIDOPHOS	41867201 148452 43197901	RBC	F	average	1.00	1.00	1.00	0.09	0.07	1.00
				biweekly	1.00	1.00	1.00	0.08	0.06	1.00
PHOSALONE	44852504 44801002	RBC	F	average	0.044	0.015	0.13	1.45	0.77	0.062
				biweekly	0.048	0.017	0.14	1.31	0.68	0.061
DIAZINON	43543901 43543902 40815003 41942002	RBC	M	average	0.12	0.024	0.63	0.40	0.22	0.18
				biweekly	0.14	0.027	0.68	0.34	0.18	0.18
DIMETHOATE	43128201 164177	RBC	M	average	0.27	0.15	0.48	0.36	0.20	0.19
				biweekly	0.25	0.13	0.47	0.40	0.22	0.15
METHAMIDOPHOS	41867201 148452 43197901	RBC	M	average	1.00	1.00	1.00	0.07	0.05	1.00
				biweekly	1.00	1.00	1.00	0.06	0.05	1.00
PHOSALONE	44852504 44801002	RBC	M	average	0.054	0.022	0.13	18.07	9.81	0.004
				biweekly	0.072	0.032	0.16	2.72	1.40	0.023

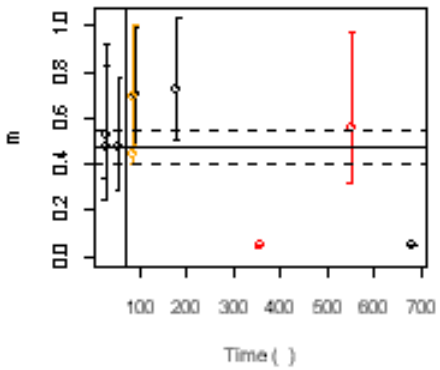
Figure III.B.3-1a. Plots of potency versus time for brain cholinesterase measured in rats exposed to diazinon

Average Dose

Female

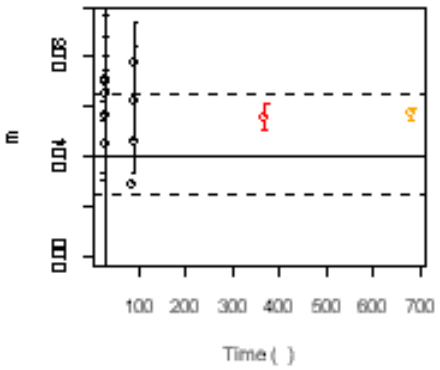


Male



Duration Specific Dose

Female



Male

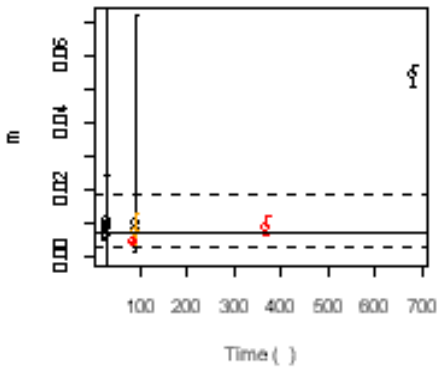
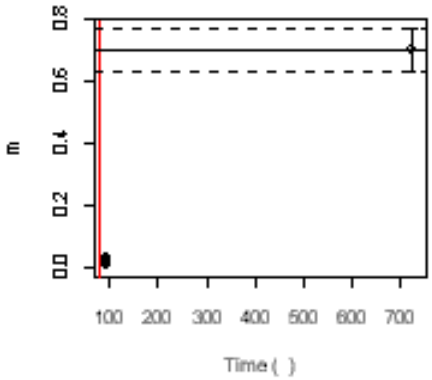


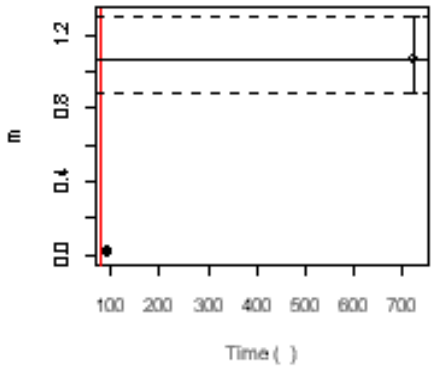
Figure III.B.3-1b. Plots of potency versus time for brain cholinesterase measured in rats exposed to dimethoate

Average Dose

Female

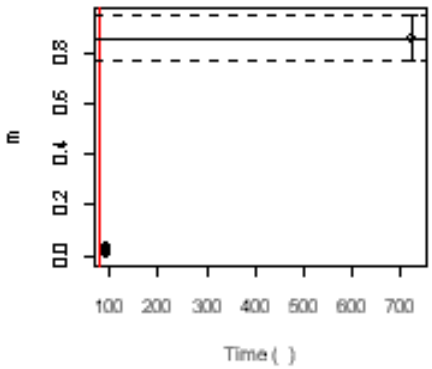


Male



Duration Specific Dose

Female



Male

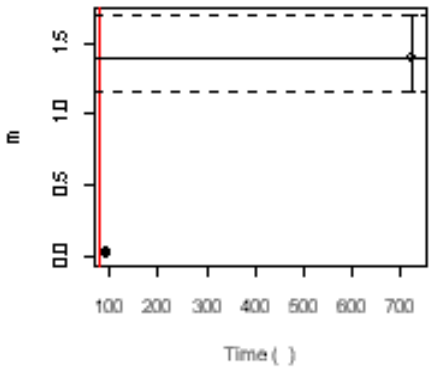
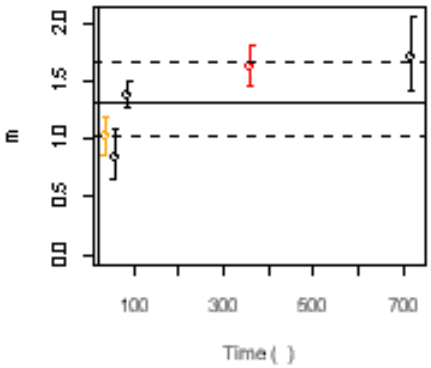


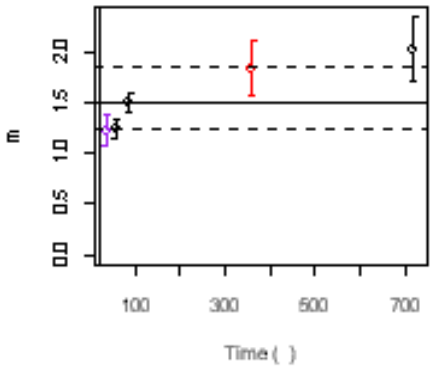
Figure III.B.3-1c. Plots of potency versus time for brain cholinesterase measured in rats exposed to methamidophos

Average Dose

Female

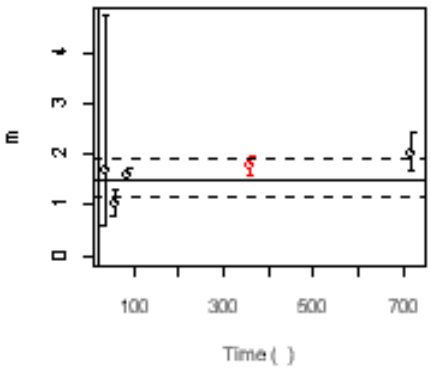


Male



Duration Specific Dose

Female



Male

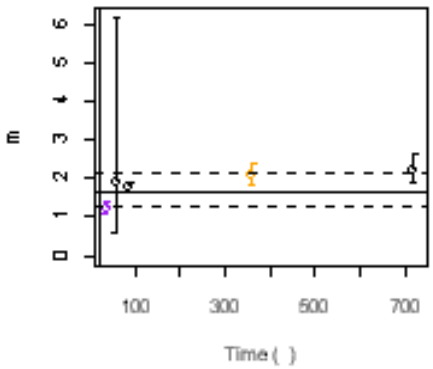
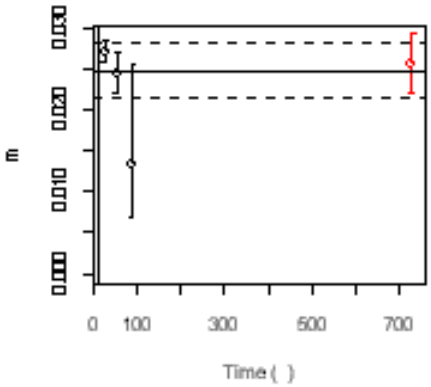


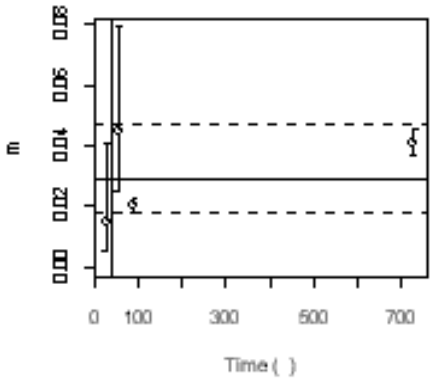
Figure III.B.3-1d. Plots of potency versus time for brain cholinesterase measured in rats exposed to phosalone

Average Dose

Female

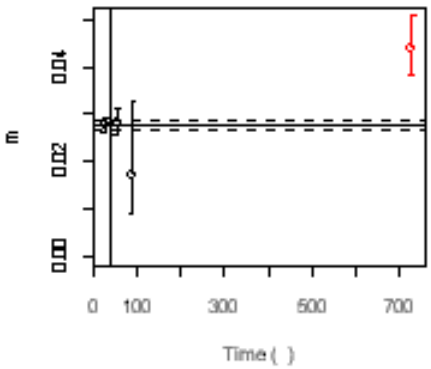


Male



Duration Specific Dose

Female



Male

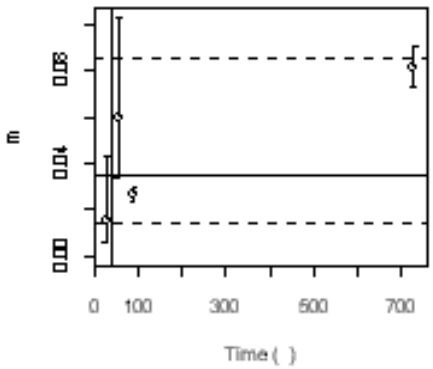
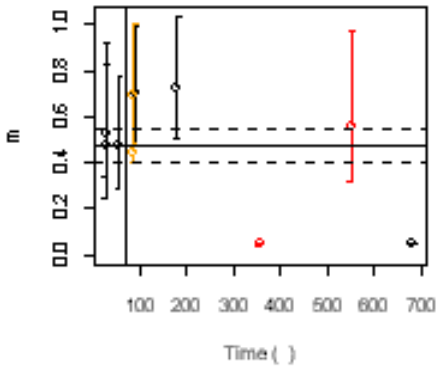


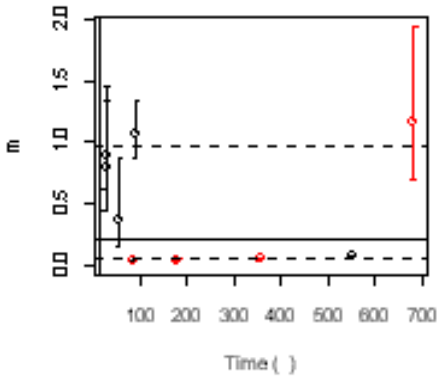
Figure III.B.3-2a. Plots of potency versus time for RBC cholinesterase measured in rats exposed to diazinon

Average Dose

Female

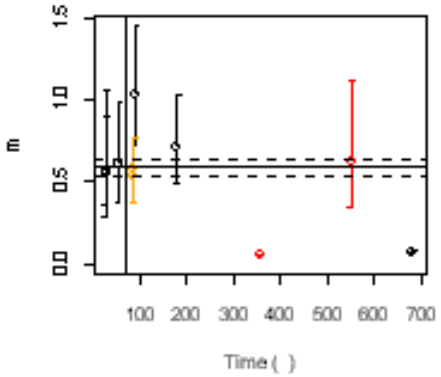


Male



Duration Specific Dose

Female



Male

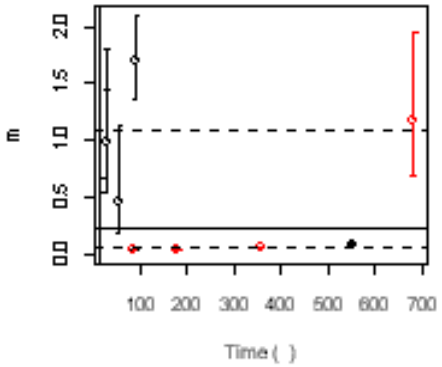
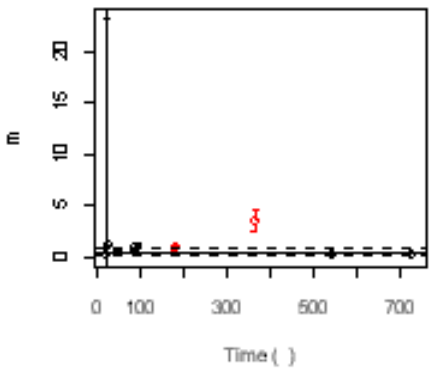


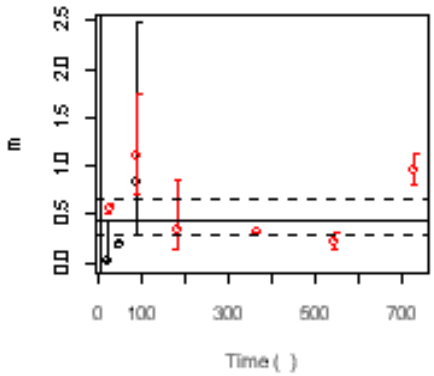
Figure III.B.3-2b. Plots of potency versus time for RBC cholinesterase measured in rats exposed to dimethoate

Average Dose

Female

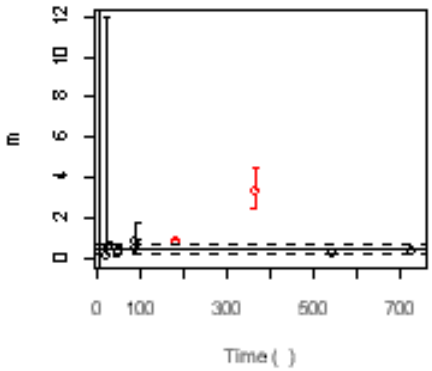


Male



Duration Specific Dose

Female



Male

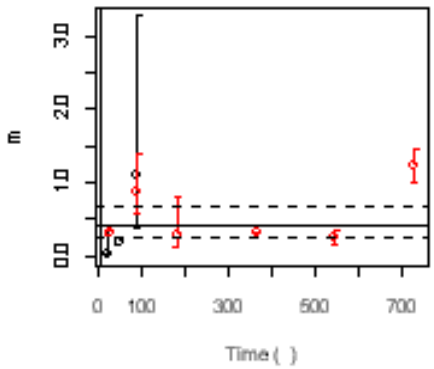
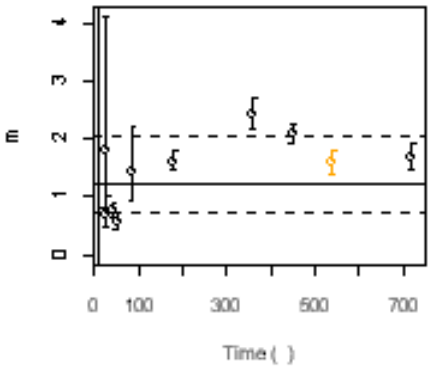


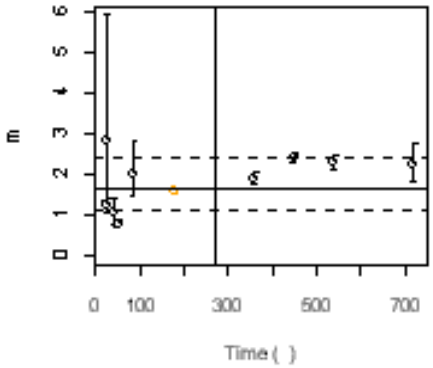
Figure III.B.3-2c. Plots of potency versus time for RBC cholinesterase measured in rats exposed to methamidophos

Average Dose

Female

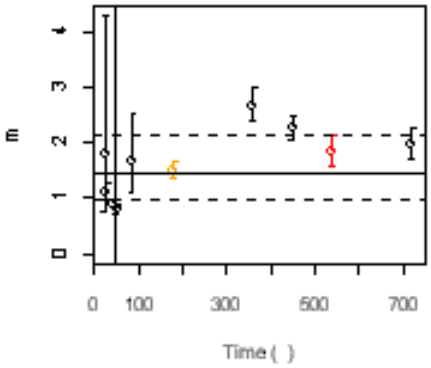


Male



Duration Specific Dose

Female



Male

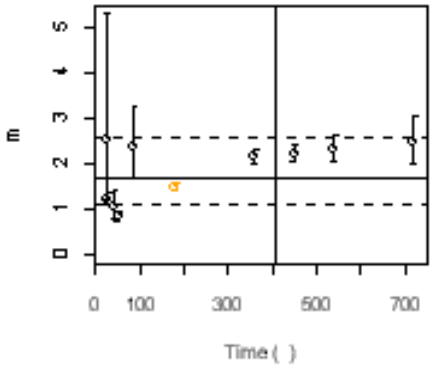
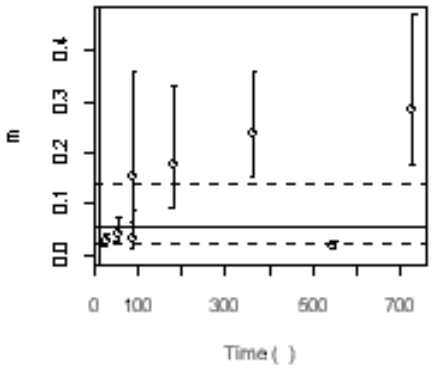


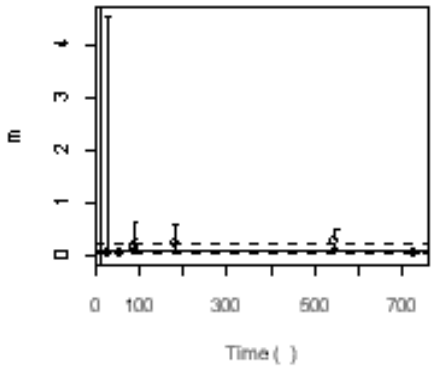
Figure III.B.3-2d. Plots of potency versus time for RBC cholinesterase measured in rats exposed to phosalone

Average Dose

Female

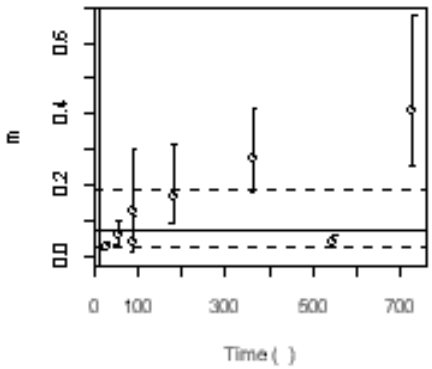


Male

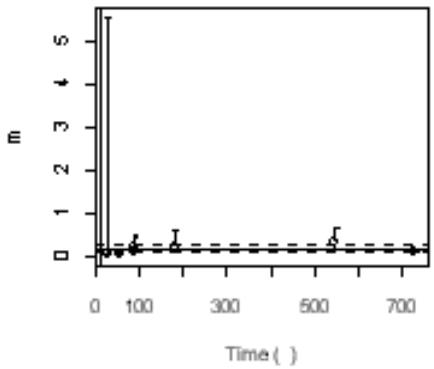


Duration Specific Dose

Female



Male



Stage 3: Compare the BMD_{10} 's and BMDL's of the index chemical calculated from the average compound intakes and the duration-specific compound intakes (Table III.B.3-3).

As shown in Table III.B.3-3, BMD_{10} and BMDL calculated using the average compound intake from July analysis are similar to but slightly smaller those calculated with the July methods with duration-specific compound intakes. BMD_{10} and BMDL calculated using the average compound intake from July analysis are similar those calculated with the December methods with duration-specific compound intakes.

Table III.B.3-3. Comparison of Average Intake *vs* Duration-Specific Intake BMD_{10}s and BMDLs

Compartment Sex	JULY				DECEMBER		
	Average Intake		Duration-Specific Intake		Compartment Sex	BMD_{10}	BMDL
	BMD_{10}	BMDL	BMD_{10}	BMDL			
FEMALE RBC	0.09	0.07	0.08	0.06	FEMALE brain	0.08	0.07
FEMALE brain	0.09	0.08	0.08	0.07			
MALE RBC	0.07	0.05	0.06	0.05	MALE brain	0.07	0.06
MALE brain	0.08	0.07	0.07	0.06			

Conclusions: The pilot analysis of compound intakes using duration specific values showed that relative potency estimates calculated from slope-scaling factors and BMD_{10}s are similar to those calculated using the average study compound intake. Based on this analysis, it is reasonable for OPP to continue using the average compound intake for its potency estimates. Concerning the PODs for the index chemical, although the values are very similar, the PODs calculated from duration-specific intake values result in slightly smaller BMD_{10}s.

b. Response to SAP Comments from March 2002

The following analyses were performed following discussion and recommendations from the February 5-8. 2002 meeting the FIFRA SAP meeting on the "Methods Used to Conduct a Preliminary Cumulative Risk Assessment for Organophosphate Pesticides":

i. Selecting the Benchmark Response Level

At the February 5-8, 2002 meeting of the FIFRA SAP, some panel members and some Public Commenters discussed the Agency's selection

of the BMD_{10} as the benchmark response level. In response to this discussion, the Agency analyzed the detection limits of the studies assessing female brain cholinesterase levels used in the Preliminary Cumulative Risk Assessment of the OPs. This analysis has shown that generally these studies can reliably detect around 10% cholinesterase inhibition and that such levels were generally achieved in the studies. *Therefore, the Agency's use of the BMD_{10} as the benchmark response is appropriate.*

According the Agency's draft benchmark dose guidance (USEPA, 2000a), generally, the response level selected to calculate the benchmark dose should lie in the low end of the range of the responses but within assay detectability. Figure III.B.3-3 shows a plot of the range of mean brain cholinesterase inhibition observed in all treatment groups (i.e., controls were not included). That figure shows that all chemicals include at least one dose level that yields approximately 10% inhibition. Thus, it is possible to directly assess the fit of the model to data in this critical region.

The ability of a study to detect a given amount of change is measured by the power of the study. In general, the power of a study depends on the sample size and the variability of the observations, measured as the standard deviation among individual measurements. Both of these factors vary among datasets in this risk assessment. The power for each study to detect a difference between control and a single treatment group of mean brain cholinesterase activity by 1%, 5%, 7.5%, 10%, 15%, and 20% has been calculated. In Figure III.B.3-4, the proportion of datasets with at least *x* power is plotted against *x* for effect levels ranging from 1-20% inhibition, and the median power (that is, the power level such that half the datasets have greater than that level of power) among those data sets to detect each change is indicated on the axis. Only at the level of a 10% change is the median power greater than 0.80, which has been a conventional goal in designing experiments. Thus, a 10% change in mean cholinesterase activity is indeed in the low end of detectability of assays for brain cholinesterase activity as they were conducted in the studies used in this risk assessment.

Figure III.B.3-3. Observed levels of inhibition relative to concurrent control for all dose-groups. The solid vertical line indicates 10% inhbition.

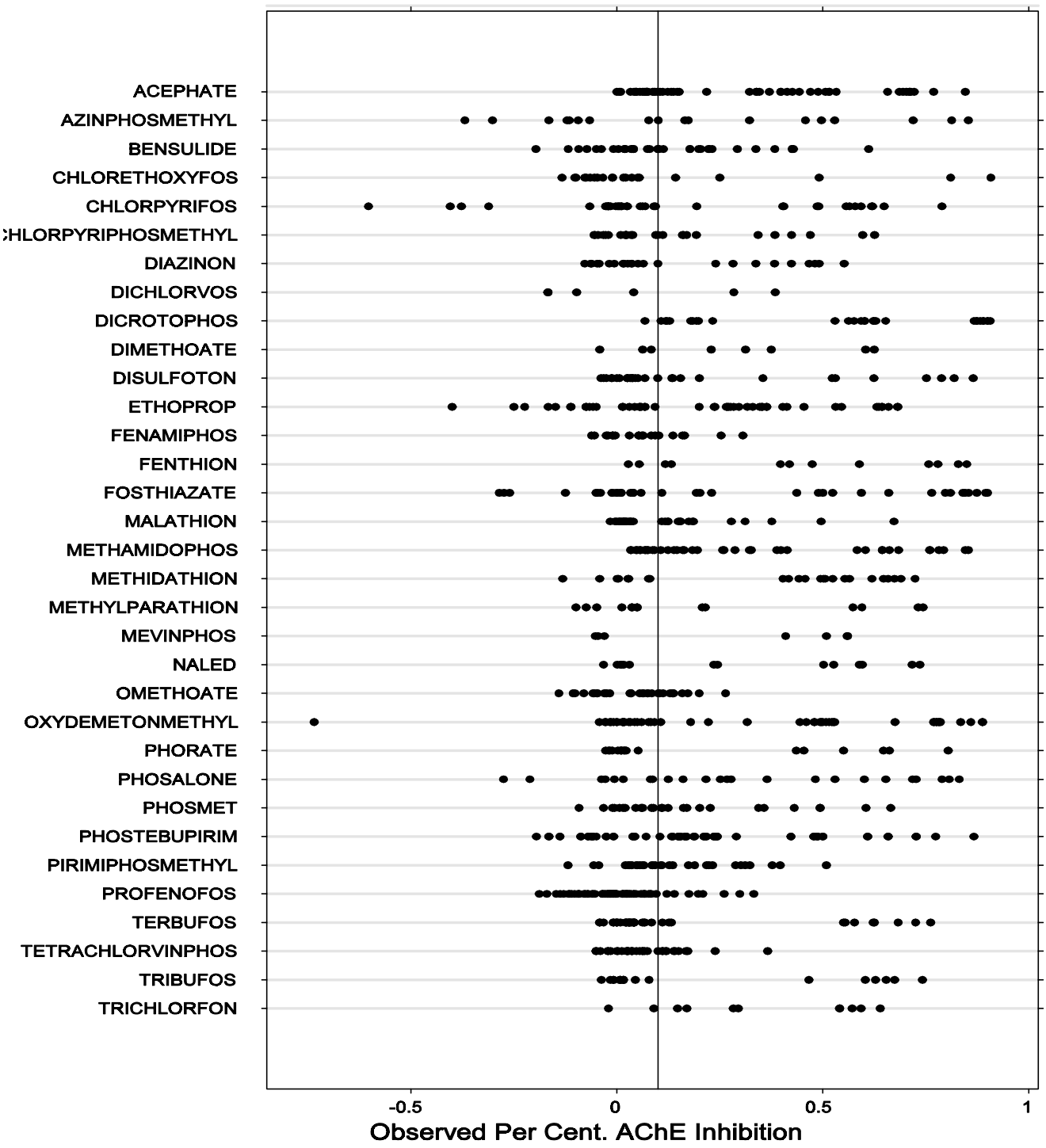
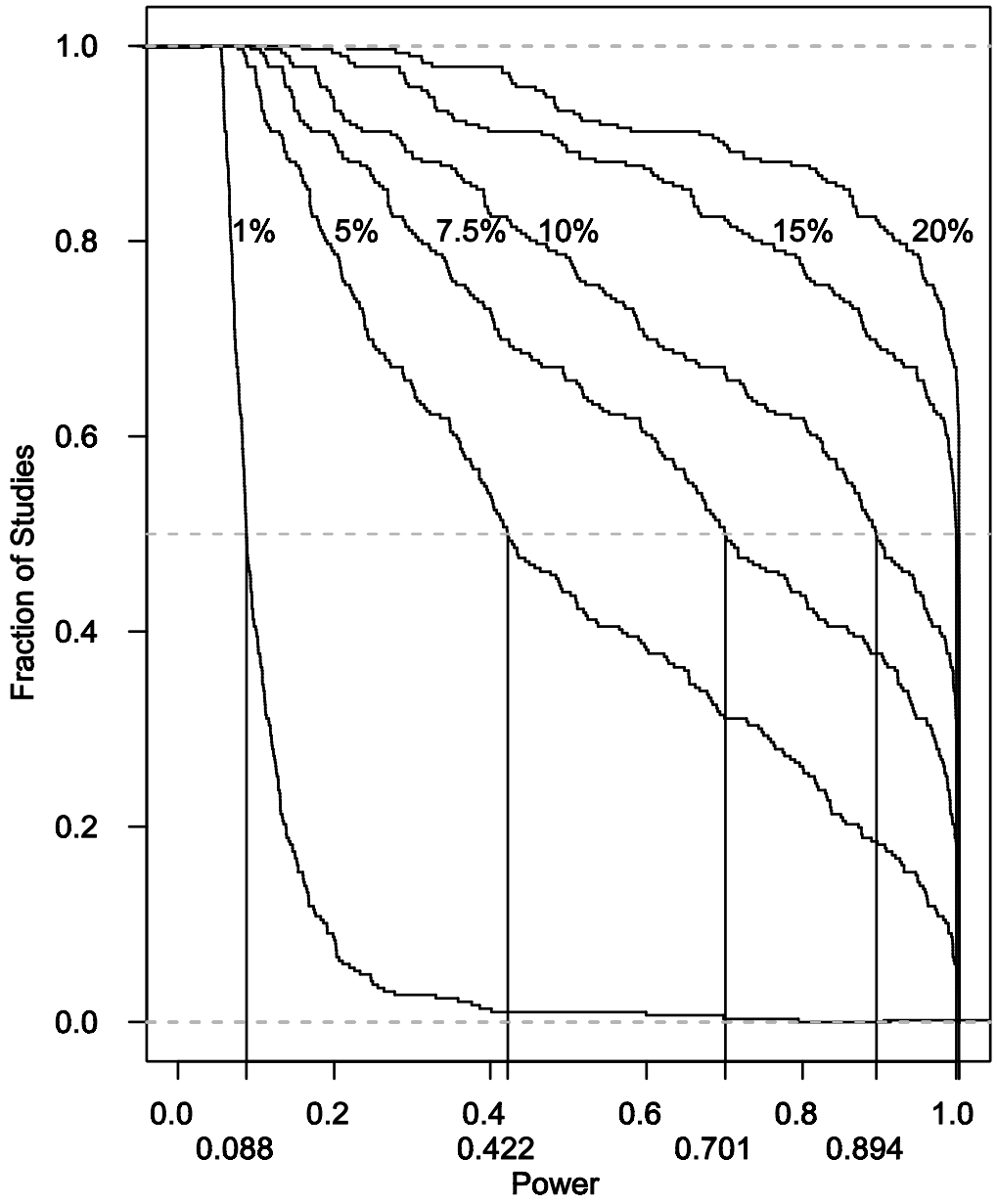


Figure III.B.3-4. Distribution of the power to detect a 1%, 5%, 7.5%, 10%, 15%, and 20% change in mean cholinesterase activity among datasets in the risk assessment. For each effect of treatment, the curves represent the fraction of datasets for which the power is at least the value on the x-axis to detect that effect. For example, half the studies have at least a power of 0.894 to detect a 10% change in mean cholinesterase activity.



ii. Standard and formal definition of the full mathematical exponential model

A formal presentation of the exponential model is included in the Appendix III.B.1.

iii. Individual Animal Data: Consequences of Aggregating Data

At the February 5-8, 2002 meeting of the FIFRA SAP, some members of the panel discussed the fact that the dose-response modeling of cholinesterase inhibition was based solely on dose group means, standard deviations, and sample sizes. The discussion centered about the issue: to what extent would the results of the analysis have differed if individual animal data had been used? The answer to that question has two parts.

1. The statistical methods used in the analysis depend on the data only through their dose group means, standard deviations, and sample sizes.

Thus, applying the same analysis to individual animal data would result in the same numerical estimates as the current analysis. The following argument shows why this is so. Whether the model fit uses generalized least squares or is a nonlinear mixed effects model (See III.B.1), the parameter estimates are the result of optimizing expressions that depend on the individual data through quadratic forms like:

$$(\mathbf{y} - \boldsymbol{\mu})' \mathbf{V} (\mathbf{y} - \boldsymbol{\mu})$$

Here, **y** is a column vector of the individual observations $\{y_{ij}\}$, *i* indexes dose group (in this discussion, "dose group" refers to the observations on animals of the same sex exposed at the same time and dose to the same chemical) and *j* indexes individual within that dose group. The vector **μ** is the vector of fitted values. Since all individuals in the same dose group were exposed to the same dose, the fitted values for each individual in a dose group are all identical. Finally, the matrix **V** is symmetric, and has the form **D** + **M**, where **D** is diagonal, and partitioned such that the values corresponding to the same dose group are identical to each other. **M** is symmetric and partitioned into blocks that correspond to the dose groups. The values within any given block are identical to each other. The partitioning of the components of **V** is due to the fact that all the individuals of the same sex given the same dose in the same study are treated identically by the model. A direct consequence of the partitioning of **μ**

and **V** is that the value of the above quadratic form can be expressed solely in terms of group means, standard deviations, and sample sizes.

2. Distribution of the brain cholinesterase data.

The methods used in the dose-response analysis assume the data is normally distributed. If the individual cholinesterase activity measurements were distinctly non-normal, it would be of interest to determine the impact of transformed or trimmed data on the benchmark dose estimates used to estimate relative potency.

Individual animal data for female and male rat brain cholinesterase activity were available for a small subset of the studies used in the the Draft Revised Cumulative Risk Assessment for the OPs. Individual animal data were available from 15 studies representing 11 chemicals (see Table III.B.3-4). Each study included several dose-response data sets in both males and females; each dose-response data set included several dose groups. (Note to the reader: Individual animal data form male and female brain cholinesterase activity used in the following analysis have **NOT** been released to the public).

i. Test for normality.

Each individual dose group (sample sizes ranging from about 5 to 50) was tested for deviations from normality using the Shapiro-Wilk test for normality (Shapiro and Wilk, 1965). The P-values for each dose group in a study were then combined using Fisher's method (Sokal and Rohlf, 1981; section 18.1), giving an overall P-value for deviation from normality for each MRID. Table III.B.3-4 gives the results of this initial test for normality.

The result of combining all the P-values over all studies was highly significant: the P-value is 9×10^{-8}. Thus, there is evidence of some deviation from normality, though, given the amount of data available for the test, and the relatively few chemicals for which the overall P-value is significant (only 2/15 MRIDs have a significant deviation from normality), the overall deviation from normality does not seem excessive

ii. Identify the nature of the deviations from normality.

Two possibilities were explored: that the data were such that a power transformation (in the form of the Box-Cox transformation; Sokal and Rohlf, 1981, section 13.9) would result in a normal distribution, and that the data were "contaminated", that is, the bulk of the observations are from a normal distribution, with an

occasional too large or too small value (Rosenberger and Gasko, 1983). The approach taken in this analysis was to use maximum likelihood to estimate the parameters in two models:

1). An observation *y* is sampled from a normal distribution with mean μ and standard deviation σ with probabilty *p*, and from a normal distribution with mean μ and standard deviation *a*×σ, where *a* > 1, with probability 1 - *p*. Here the mean and standard deviation are specific to each dose group, but *a* is the same value for all dose groups in a study.

2) If the data *y* were transformed to *z* by the Box-Cox transformation: $z = \frac{y^t - 1}{t}$ (if $t \neq 0$) or $z = \log(y)$ (if $t = 0$), the transformed data would be normally distributed, with separate mean and standard deviation for each dose group (but only one power parameter *t* for each study). When *t* = 1, then *z* = *y* - 1, and the original variable *y* is normally distributed.

The Akaike Information Coefficient (AIC; Burnham and Anderson, 1998) was calculated for each of the two hypothetical distributions for each study. AIC is useful for comparing different probability models fit to the same data sets: smaller AIC values indicate better fits. Table III.B.3-5 shows the AIC values that resulted from fitting the two models just described to the individual animal data from each study. In addition, the power parameter estimated in the Box-Cox model was tested for significant difference from one.

For eight of the fifteen studies, the AIC for the Box-Cox transformed data was less than that for the contaminated normal. Only two of those studies had a Box-Cox parameter significantly different from one, indicating that a Box-Cox transformation would result in a significantly more normal distribution. In the remaining seven of the fifteen studies, including the two with significant Shapiro-Wilk tests, the contaminated normal model provides a better description of the data. The overall AIC for the contaminated normal distribution is less than that for the Box-Cox transformed data, showing that the contaminated normal model is superior to the Box-Cox model as a single overall probability model for these data.

iii. Impact of non-normality on the BMD estimates.

BMD_{10}s were calculated for trimmed and untrimmed data. Table III.B.3-6 shows the results of applying the Shapiro-Wilk test to the trimmed individual data. The overall P-value for all the data taken together is 0.056, indicating a substantial improvement

Aggregated datasets were produced from the original (untrimmed) individual data and the trimmed individual data, and both the basic and expanded models fit to each set of data for each chemical (See I.B and III.B.1). Four chemicals were affected by the trimming: dicrotophos, methamidophos, phorate, and phosalone. Thus, comparisons between untrimmed and trimmed data is limited to nine studies from four OPs.

Table III.B.3-7 compares the BMD_{10} calculated from the original data to that calculated using the trimmed data, for both basic and expanded models. The largest difference is less than 20% of the untrimmed value, which is reasonably small. *The current dose-response analysis used in the Draft Revised Cumulative Risk Assessment of the OPs, based solely on aggregated data, is relatively robust to the kinds of deviations from normality identified here.*

In summary, since the statistical methods used to fit dose-response models to the data depend on the data only through their means, standard deviations, and sample sizes, the only way an analysis of individual data might differ from that of aggregated data would be if the distribution of the data were substantially non-normal. The distributions of a subset of the data were examined, resulting in evidence that some studies did produce data that deviated from normality. When extreme observations were omitted, the overall distribution of the data became closer to a normal distribution. However, benchmark doses calculated using the trimmed data, were quite similar. to those using all the data. *Thus, it is unlikely that using aggregated data has substantially distorted the estimates of benchmark doses that would obtain had the analysis been based on individual animal data.*

Table III.B.3-4. Chemicals and studies used in individual animal analysis.

Chemical	Study (MRID no.)	Number of Dose Groups	Number Failed	Proportion Failed	Combined Shapiro-Wilks P-value
Methamidophos	148452	20	8	0.400	1.63e-08
Methamidophos	41867201	20	2	0.100	1.08e-01
Methamidophos	43197901	8	1	0.125	1.37e-01
Fenamiphos	44051401	8	1	0.125	1.60e-01
Bensulide	44161101	32	4	0.125	8.14e-02
ODM	44189501	36	1	0.028	8.09e-01
Fosthiazate	44269905	14	1	0.071	3.23e-01
Dicrotophos	44527802	16	4	0.250	1.01e-03
Phosalone	44801002	8	1	0.125	6.52e-02
Phosmet	44811801	16	1	0.063	4.63e-01
Terbufos	44842302	8	1	0.125	8.98e-02
Phosalone	44852504	24	0	0.000	2.60e-01
Phorate	44895301	8	1	0.125	7.55e-02
Phorate	44895302	10	0	0.000	2.71e-01
Chlorpyrifos-methyl	44906902	10	1	0.100	5.47e-02

"Number of Groups" is the total number of dose groups available; "Number Failed" is the number of individual dose groups for which the Shapiro-Wilks test reported a P-value less than 0.05; "Proportion Failed" is the proportion of dose groups that failed the test (Number Failed/Number of Groups); "Combined Shapiro-Wilks P-value" is the overall P-value for each MRID, resulting from using Fisher's method to combine the P-values for the individual dose-group tests.

Table III.B.3-5. AIC values for the Box-Cox and the contaminated normal models.

Chemical	Study (MRID no.)	AIC	
		Contaminated Normal	Box Cox Transformed
Methamidophos	**148452**	***362.46***	386.90
Methamidophos	41867201	***250.84***	251.01
Methamidophos	43197901	64.15	***62.26***
Fenamiphos	44051401	***105.23***	106.66
Bensulide	44161101	5586.60	***5578.49***
ODM	44189501	527.11	***524.59***
Fosthiazate	44269905	2242.37	***2239.92***
Dicrotophos	**44527802**	***213.37***	228.71
Phosalone	44801002	174.91	***164.19 ****
Phosmet	44811801	382.29	***380.07***
Terbufos	44842302	***339.71***	340.77
Phosalone	44852504	***204.14***	209.10
Phorate	44895301	***108.88***	119.00
Phorate	44895302	366.05	***359.44 ****
Chlorpyrifos-methyl	44906902	211.33	***208.39***
Sum:		**11139.44**	**11159.50**

MRID numbers for data that were significantly non-normal by the Shapiro-Wilks test (see Table III.B.3-4) are written in bold. The smaller of the two AIC values for each MRID is written in bold italics. When the Box-Cox power parameter is significantly different from 1, the Box-Cox AIC is followed by an asterisk.

Table III.B.3-6. P-values for the Shapiro-Wilks test, combined over all dose groups in a study for the trimmed individual data.

Chemical	Study (MRID no.)	P.value
Methamidophos	148452	0.046
Methamidophos	41867201	0.293
Methamidophos	43197901	0.137
Fenamiphos	44051401	0.160
Bensulide	44161101	0.081
ODM	44189501	0.809
Fosthiazate	44269905	0.323
Dicrotophos	44527802	0.937
Phosalone	44801002	0.065
Phosmet	44811801	0.463
Terbufos	44842302	0.090
Phosalone	44852504	0.863
Phorate	44895301	0.831
Phorate	44895302	0.271

Table III.B.3-7. Benchmark doses from the basic and expanded models for untrimmed (original) and trimmed data.

Chemical	Data Treatment	Female BMD_{10}	
		Expanded Model	Basic Model
Dicrotophos	original	NA	0.032
	trimmed	NA	0.026
Methamidophos	original	NA	0.080
	trimmed	NA	0.079
Phorate	original	0.215	0.036
	trimmed	0.201	0.037
Phosalone	original	6.426	3.843
	trimmed	6.313	3.847

NA: As shown in I.B, the basic model was used to estimate potency for methamidophos and dicrotophos.

References

Burnham, K. P. and Anderson, D. R. 1998. *Model Selection and Inference. A Practical information-Theoretic Approach.* Springer. New York.

Rosenberger, J. L. and Gasko, M. 1983. Comparing location estimators: trimmed means, medians, and trimean. Chapter 10 in *Understanding Robust and Exploratory Data Analysis*, David C. Hoaglin, Frederick Mosteller, and John W. Tukey, eds. Wiley. New York.

Shapiro, S. S. and Wilk, M. B. (1965). "An analysis of variance test for normality (complete samples)", Biometrika, 52: 591-611.

Sokal, R. R. and Rohlf, F. James. (1981). *Biometry, Second Edition.* Freeman. San Francisco.