III. Appendices

B. Hazard/RPF

1. Technical Aspects of Dose-Response Analysis

a. Analysis of Individual Data Sets (July 2001)

i. Summary of July 2001 Analysis

The goal of the statistical methods was to estimate a quantity that would be proportional to the potency of each chemical. along with confidence intervals. The data for this study were in the form of dose-response studies which measured the effect of different concentrations of OP pesticides on cholinesterase activities in brain, red blood cells, and plasma. The mean and standard deviation of cholinesterase activity, and number of animals examined were available for several dosages in each data set. Females and males were analyzed separately in each study. Studies were nested: for each chemical there were several groups of studies, each with a different MRID; within MRID, one or more studies were conducted, each with measurements taken for several durations of exposure. It is possible that potency increases to an asymptotic value as exposure duration increases. Studies with the same MRID were conducted in the same laboratory. Thus several steps were required to analyze the collection of data sets for each compartment × sex combination for each OP pesticide:

1) Adequately model the relationship between dose and cholinesterase activity for each individual study, and estimate the absolute potency for that study.

2) Determine which exposure durations are likely to be long enough that potencies are close to the steady-state value

3) Combine potency estimates within MRID, resulting in a single estimate for each MRID, with standard error.

4) Combine potency estimates across MRIDs, resulting in a single potency estimate for the chemical × compartment × sex combination

5) Compute the relative potency by dividing all the potency measures within sex and compartment by that of the index compound, and estimate the standard error of the result.

6) Compute BMD and BMDLs for each data set for the index compound, and combine the estimates using the same methods as for potency.

The following sections describe each of these steps in greater detail.

ii. Dose-Response Modeling (July 2001)

The cumulative dose-response assessment for this analysis is based on the Relative Potency Factor (RPF) methodology (U.S. EPA, 2001). The RPF approach assumes that the dose-response for a combination of exposures, at least for relatively low levels of exposure, are dose-additive. That is, the response *y* for a combination of exposures D_i would be :

$$y = f\left(\sum_i m_i D_i ; \beta\right) \qquad \textbf{Eqn III.B.1-1}$$

where m_i is the absolute potency of the i^{th} exposure, and β is a vector of parameters of the dose-response function whose values are the same for all chemicals. In practice, we select one index chemical, (call it *I*), and express all potencies as ratios to that of the index chemical; $R_i = m_i/m_I$. Then the expected response *y* for a combination of exposures D_i would be:

$$y = f\left(m_I \sum_i R_i D_i ; \beta\right) \qquad \textbf{Eqn III.B.1-2}$$

The response is the same as it would be if a dose of the index compound equal to $\sum_i R_i D_i$ had been given. Furthermore, any combination of dosages that give the same overall weighted sum should result in the same response. For example, suppose the following potencies apply to four chemicals:

Chemical	Potency $(mg/kg/day)^{-1}$ (m_i)	Relative Potency (R_i)
A	0.346	0.309
B	0.0082	0.0073
C	1.21	1.08
D [index]	1.12	1.00

Suppose two sets of doses:

Chemical	Set 1	Set 2
A	0.0137	0.00324
B	1.17	0.00162
C	0.0391	0.0486
D	0.0391	0.0405
$\sum D_i$	1.26	0.094
$\sum R_i D_i$	0.094	0.094

Even though the two sets of doses are very different, and result in quite different total doses (1.26 mg/kg/day in set 1, 0.094 mg/kg/day in set 2), the expected responses are identical. If $BMDL_l$ is the lower confidence limit for a specified response (the benchmark response, BMR, in this case, a 10% reduction in cholinesterase activity) for the index chemical, then the response to the combination of exposures represented by the D_i is likely to be smaller than the BMR if $\sum R_i D_i$ is smaller than $BMDL_l$.

In this analysis, the dose-response function had to accommodate two important features of the data. First, since the results of multiple studies, perhaps carried out in different laboratories and at different times, and even sometimes reporting AChE activities in different units, it seemed prudent to express activity at a given dosage as a fraction of control activity. Implicit in this formulation is the idea that the among-data-set component of variability follows a multiplicative error distribution. Second, it was observed that, as doses increased, AChE activity in quite a few data sets approached a lower non-zero asymptote. This asymptote varied among data sets, chemicals, and compartments. These two properties of the data were accommodated by fitting the model:

$$y = B + (A - B)\mathbf{e}^{-mD} \qquad \textbf{Eqn III.B.1-3}$$

where *A* is the background level of cholinesterase activity, and *B* is the limit of AChE activity for large doses. In practice, in some cases, it was not possible to estimate all three parameters for a data set, or this model failed to adequately fit the data. In these cases, *B* was often set to 0 and higher doses were dropped from consideration, as will be described below in more detail. To force all parameter estimates to be non-negative, what was actually estimated in each case was the natural logarithm of the parameter. So, for example, the parameters estimated were *lA, lB,* and *lm*, with $A = e^{lA}$, for example. Standard errors of the log-parameters reported by the statistical software were transformed using the delta

method to be on the appropriate scale. Parameters for this model for each data set were estimated using generalized nonlinear least squares (GNLS). GNLS was selected because it does not require distributional assumptions about the individual data (which could not be checked, since only summary statistics were available), unlike maximum likelihood (ML) estimation, and it is more robust than ML (Davidian and Giltinan, 1995, pp. 31, 39, 59). GNLS is an extension of weighted nonlinear least squares. In weighted nonlinear least squares, the parameter estimates are the values for the components of β (that is, in this case, the vector [*A, B, m*]′) that minimize the weighted sum of squares $SS = \sum w_i \left(y_i - f\left(Dose_i ; \beta \right) \right)^2$, where y_i is the observed activity, *f()* is the function described above in eq. I-3, and the weights w_i are already known and are proportional to the reciprocal of the variance. In GNLS, the weights are taken to be a known function of the mean. In the OP data for this analysis, the variance among observations within dose groups is approximately proportional to the square of the mean of the group (see fig [1]), so regression weights based on the square of the estimated mean were used to improve the efficiency of the estimates over what would be obtained with unweighted regression.

Goodness of fit of each fitted model to the corresponding data was quantified through a global test of goodness-of-fit, specifically the Pearson chi-squared statistic, through visual inspection of graphs, and through examining tables of standardized residuals. The Pearson chi-squared statistic is $\mathrm{X}^2 = \sum_i \left(\bar{y}_i - f\left(Dose_i ; \hat{\beta} \right) \right)^2 / \left(\hat{\sigma}^2 / n_i \right)$, where *i* indexes dose groups. If the model is true, then X^2 will be distributed (approximately) as Chi-squared with degrees of freedom equal to the number of dose groups minus the number of parameters estimated.

The process for getting the final parameter estimates for a data set was as follows:

1) Estimate A, B, and m using GNLS for all dose groups in the dataset.

2) If the P-value for the X^2 statistic was greater than 0.05, then the result is the estimate used. Otherwise (that is, if the P-value was less than 0.05, or no estimates resulted because the model did not converge), set B to zero, and try again with all the data.

3) If the P-value is still less than 0.05, or there is no model fit at all, then sequentially drop the remaining highest dose and refit the model with B set to zero until either the P-value exceeds 0.05, or there are only three doses remaining.

iii. Identifying Steady State

The sets of data for each chemical, sex, and compartment included a range of exposure durations. To determine which data sets had a sufficiently long exposure duration that potency was no longer changing with time, we regressed the estimated potency against exposure duration, weighting observations by the reciprocal of the squared estimated standard errors. The data for the shortest remaining exposure duration in the data set was repeatedly removed until a data set was derived in which the slope of potency versus time was not significant (that is, the P-value exceeded 0.05). In any case, the process was stopped when only three distinct durations remained. After a first pass through the data, a single duration was identified such that exposures exceeding that duration rarely showed a significant increase with time; all exposures less than that duration were removed from further consideration.

iv. Combining Potency Estimates and Computing Relative Potencies

Potency estimates were nested in two levels for each chemical × sex × compartment: generally several data sets, representing a range of exposure durations and some duplication within each MRID, and several MRIDs. Since the data sets representing exposure durations at which steady state had not been achieved were deleted from the study before this stage, it was reasonable to model the individual potency estimates as coming from a nested hierarchical sampling scheme:

1) First, assume there is an overall mean potency for a given chemical × sex × compartment combination; the procedure described below is designed to estimate this quantity, which will be used for computing the relative potency.

2) Because of small differences in husbandry, analytic procedures, and other laboratory procedures, the potency realized in MRIDs may vary among MRIDs. Model this as sampling a MRID-specific potency from a distribution centered about the overall mean potency, with variance σ^2_{MRID}.

3) Again, because of differences among the studies that contributed the data sets within a given MRID, potencies realized within a given data set may vary among data sets within an MRID. Model this as sampling a data-set- specific potency from a distribution centered about the MRID-specific potency, with variance σ^2_{DS}.

4) Finally, because a study is based on only a finite sample of animals, we can only estimate the data-set specific potency, with an error variance of σ^2.

Our final goal is to estimate the overall mean potency for the chemical × sex × compartment combination, along with a standard error that reflects the uncertainty in the estimate due both to errors in estimation and the variances among the MRID-specific and data-set-specific potencies. Furthermore, once each such overall mean potency is computed and an index chemical selected, we want to compute the potency for each chemical relative to the index chemical in each sex × compartment combination.

To facilitate this latter computation, all operations were conducted with logarithms of the potencies. Thus, in the end, the relative potencies were computed as $e^{logm-logmI}$, where "logm" is the logarithm of the potency for the chemical for which the relative potency is being calculated, and "logmI" is the logarithm of the potency of the index chemical. The uncertainty of the relative potency is then expressed as a confidence interval, obtained by exponentiating the endpoints of the confidence interval for the difference "logm - logmI", whose standard error is $\sqrt{se^2_{\log m}+se^2_{\log mI}}$, where se^2_Q is the square of the standard error for parameter *Q*, here one of the log potencies.

The estimates of logm were constructed in two stages: first, estimates of data-set-specific logms were combined, using their estimated standard errors, to yield MRID-specific estimates of logm and standard errors; then, via the same process, the MRID-specific estimates of logm were combined to yield overall mean estimates. The procedure used to combine the estimates is known as the "global two-stage method" (Davidian and Giltinan, 1995, pp 138ff). The logic of the global two-stage method is simple, and is illustrated here for estimating a MRID-specific estimate of logm. Since we have individual estimates of logm for each data set, it is natural to estimate a MRID-specific estimate as the mean of the individual data-set-specific estimates. However, if we estimate the standard error in the usual way, based on the standard deviation of the data-set-specific estimates, it turns out that the resulting estimates of the standard error are biased upwards, because the procedure described here ignores the uncertainty of the individual estimates. The global two-stage method corrects this bias by explicitly taking into account that uncertainty.

v. Estimating BMDs and BMDLs

Benchmark doses were estimated for a 10% reduction in activity from background. That is, the benchmark dose is the value *D* such that:

$$0.1 = \frac{1}{A}\left\{A - \left[B + (A - B)\mathbf{e}^{-mD}\right]\right\} \qquad \textbf{Eqn III.B.1-4}$$

which is:

$$D = -\frac{1}{m}\ln\left(\frac{0.9A - B}{A - B}\right) \qquad \textbf{Eqn III.B.1-5}$$

The BMDL was based on the lower 95% confidence limit on the estimate of the BMD. The confidence limit was computed from the estimated standard error of the reciprocal of the BMD estimate, because simulations (described in the next subsection), indicated that such confidence limits came closer in practice to the theoretical coverage than did limits based on standard errors computed on the original scale or on the logarithm scale. That is, the lower 95% confidence limit was computed as:

$\frac{1}{BMDinv + 1.645se_{BMDinv}}$, where *BMDinv* is the reciprocal of the BMD estimate:

$$BMDinv = \frac{m}{\ln\left(\frac{A - B}{0.9A - B}\right)} = g(\beta),$$

where β is the vector of parameters, [*A*, *B*, *m*]′, and se_{BMDinv} is its estimated standard error, computed as

$$se_{BMDinv} = \left(\frac{\partial g}{\partial \hat{\beta}}\right)' \hat{\Sigma} \left(\frac{\partial g}{\partial \hat{\beta}}\right),$$

where $\hat{\Sigma}$ is the estimated covariance matrix for the parameters, and $\partial g / \partial \hat{\beta}$ is the gradient of *g()* evaluated at the parameter estimates.

vi. Simulations to Support the BMDL Computations

1) Simulation Methods

The performance of the statistical methods used in OPCumRisk, as well as the behavior of the estimation process, was checked by simulation. The simulation process was as follows:

① Generate simulated data sets

a. For a set of values of *A*, *B*, and *m*, different levels of maximum inhibition observed (that is, given *A*, *B*, and *m*, pick the dose that yields the specified maximum level of

inhibition), and particular dose regimen, compute mean cholinesterase activity levels (μ) using:

$$\mu(\text{dose}) = B + (A - B)\, e^{-m \times \text{dose}}$$

b. Compute the standard deviation of data from each dose (σ) by:

$$\sigma(\text{dose}) = CV * \mu(\text{dose})/100.$$

c. For each of 500 replicate data sets, sample means from a normal random number generator with mean given as in (a) and standard deviation given as $\sigma/\sqrt{}$(Sample Size). Sample standard deviations by generating Chi-squared random numbers with degrees of freedom equal to (Sample Size - 1) using a Chi-squared random number generator, then multiplying by the square of the required standard deviation, dividing by Sample Size, and taking the square root.

d. The parameter values used for the simulation were based on values observed in a pilot sample of studies, and are given in the table below:

level of *A*:	2000
levels of *B* (*B/A*):	0, 500, 1000 (0, 0.25, 0.5)
levels of *m*:	0.03, 0.20, 1.0, 5.0
Highest dose selected to give activity at highest dose = A - (1 - ActMaxF)*(A - B)	ActMaxF=(0.5, 0.85, 0.95)
Dose Regimens (fractions of highest dose):	{0, 0.05, 0.20, 1.0}, and {0, 0.01, 0.067, 0.3, 1.0}
CV:	20%, 40%
Sample Sizes:	6, 10

A total of 216 unique combinations of parameter values were simulated, with each represented by 500 simulated data sets, for a total of 108,000 different simulated data sets available for parameter estimation.

② Loop through the simulated data sets:

a. Attempt to fit a model, estimating B, to all data points.

b. If the P-value for the Pearson goodness-of-fit chi-square was less than 0.05, or the model did not converge, try to fit a model with B set to 0.

c. Repeat step (b), dropping the high dose group with each iteration, until the Pearson chi-square is greater than 0.05, or only three doses remain.

d. Record parameter estimates, the covariances of the estimates, the estimated benchmark dose for a 10% reduction of activity, and 95% lower confidence limits based on standard errors for (i) BMD computed on the original dose scale, (ii) the natural logarithm of BMD, and (iii) 1/BMD.

③ Summarize the simulation results. For each combination of parameter values, collect:

a. The bias [mean(*m*) - true(*m*)] and relative bias [(mean(*m*) - true(*m*))/true(*m*)] of the estimate of *m*.

b. The proportion of each of the approaches for computing confidence limits in which the calculated limit is less than the true BMD. The true BMD is computed from the values of *A*, *B* and *m* from which the data were generated.

c. Analyze the simulation results to examine relationships between parameter values and levels of bias or BMD coverage, and to determine which of the approaches to computing the BMDL comes closest to providing the nominal 95% coverage. The analysis is based largely on the use of regression trees (Breiman et al., 1984) as implemented in the Rpart package version 3.0.1 for R version 1.3.0.

The simulations showed that the bias on the absolute potency estimate in an individual data set depends upon the true value of B, and on the degree of cholinesterase inhibition at the high dose. Three different levels of B were considered in the simulations: 50% of background, 25% of background, and 0. Note that here, B refers to the value of the dose-response model used to generate the data, NOT the value estimated by fitting the data. Figure 2 shows that, if B was 50% of background, then *m* was underestimated by about 28%. None of the other factors in the simulation was systematically associated with variation around this level. If B was a smaller fraction of background, either 0 or 25% of background, then the bias depended upon how

much of the dose-response curve was captured in the individual dataset. If the cholinesterase activity at a high dose level was at least 85% of the estimated B, then on average absolute potency was over estimated by about 3%. If the cholinesterase activity at the high dose level was only 50% of the estimated B, then again, the degree of bias depended upon the true value of B. For example, if the true value of B was 25% of background, then the absolute potency was overestimated by around 11%. When B was 0, potency was overestimated by over 50%.

If the true value of B was known for a given dataset, the degree to which the estimate of absolute potency for that data set is likely to be biased could be evaluated, since both conditions that influence bias in this simulation depended upon the true value of B. However, B is unknown for real data sets. To try to evaluate the magnitude of bias in the estimates of absolute potency that might be expected, we have estimated B/A for each chemical × sex × compartment combination by combining estimates of B/A from data sets in those compartments which were adequately fit by the full model (that is, while estimating B). Also, the ratio of the model-predicted activity at the highest dose actually used in the fitting to the predicted control activity was computed for each data set (using predicted values smooths out some of the statistical fluctuations). That ratio was used to estimate ActMaxF for each data set. The degree to which potency estimates might be biased by cross-tabulating the estimate of B/A and the estimate of ActMaxF can be evaluated. Although the simulation was run at discrete values of B/A and ActMaxF, of course the real data are distributed continuously across the possible ranges of these two variables. To do the cross-tabulation, both variables were broken at the midpoints between discrete points used in the simulation: B/A at 0.125 and 0.375, and ActMaxF at 0.675. Altogether, B could be estimated for 1135 data sets. The breakdown by B/A and ActMaxF and the of the bias in the absolute potency estimate the simulations predict would be operative if every dataset in the corresponding group of the table had values of B/A and ActMaxF at the corresponding simulation value is:

	ActMaxF	
B/A	≤0.675	> 0.675
≤0.125	73 (6.4%)	117 (10.3%)
	bias: *50%*	bias: *3%*
0.125 < B/A ≤0.375	199 (17.5%)	382 (33.7%)
	bias: *11%*	bias: *3%*
0.375 < B/A ≤1	139 (12.2%)	225 (19.8%)
	bias: *-30%*	bias: *-30%*

About 61.5% of these data sets fall into the 3% or 11% bias groupings, while about 32% of the data sets fall into groupings where the simulations predict that absolute potency would be underestimated by about 30%.

Two complications in the calculation of the relative potency factors need to be considered for a total evaluation of bias. First, since the RPFs are based on average absolute potencies, data sets with both high and low values of ActMaxF would be combined to get an overall average. This should mean that bias in the overall average potency should fall somewhere between that of the individual data sets. Secondly, the denominator of the RPF, the absolute potency of methamidophos, itself must have some bias. This has the effect of reducing the overall bias of the RPF of chemicals whose absolute potencies were overestimated, and increasing the bias of the RPF of chemicals whose absolute potencies were underestimated. In fact, it is likely this is a small effect, since, in the RBC compartment in males, the ratio B/A is about 0.14, so the bias in its absolute potency would range between 3% and 11% based on the simulation.

It should be clear that the overall bias of the relative potency estimates that is due to the estimation procedure is likely to be relatively small. However, this whole analysis should be taken as suggestive, rather than determinative, of the levels of bias likely to exist in the estimates of absolute potency. The real data sets have a range of dose-placements and sample sizes, while the simulations, while based on the distribution seen in the data, used a much smaller range. We have estimated B/A, but many data sets do not allow B to be estimated. By dropping those data sets, we may have biased the estimate of B. Finally, this is not a very quantitative analysis of any bias that might result from combining data sets.

2) Summary of Simulation Results

① Bias of Potency Estimates (Figure 2). The primary determinants of the relative bias in potency estimates is the value of B and the activity at the highest dose (ActMax). In models with B = 1000 (that is, B is half of A), the potency is underestimated by about 28%. In models with smaller B, if doses are large enough that the activity at the highest dose is close to B, the bias is about 3% of the true value. The largest bias, around 54%, occurs in models with B=0, and where the activity at the highest dose is half the background.

② Coverage of Nominal 95% Confidence Interval (Figure 3). In general, confidence limits for the BMD in individual data sets have lower than their nominal coverage; that is, the BMDL is too high. Only if B=0, where the average coverage is about 97% or, for larger B, the activity at the highest dose is only half way to the horizontal asymptote, where the average coverage is about 93%, are coverages close to nominal.

vii. Software

To facilitate modeling the large number of datasets evaluated in this study, special purpose software was written using version 1.2.1 of the open source statistical programming language R, (Ihaka and Gentleman, 1996; http://cran.r-project.org). A graphical user interface using the *tcltk* package for R was constructed to facilitate all phases of the analysis. Model parameters and their standard errors were estimated using the function *gnls* in the R package *nlme* (version 3.1-10; see Pinheiro and Bates, 2000).

viii. Conclusion (July 2001)

The present approach to determining relative potency has several advantages. As opposed to another method, such as maximum-likelihood, the generalized least squares method used here for estimation of the parameters of the individual dose-response curves is generally more robust to misspecified data distributions which is important since actual data distributions were not directly available for checking. A novel aspect of this analysis was the use of a hierarchical statistical model to combine estimates of potency for the oral studies (average absolute potency values) and to combine estimates of benchmark dose for the oral, dermal, and inhalation routes (average BMD_{10}s for the index chemical). Historically, OPP has selected single data sets or data points (such as reference doses [RfD] or NOAELs) for use in single chemical risk assessment. Aggregating over multiple data sets from studies with relatively well-defined study design has the advantage of being able 1) to

increase the precision of the estimates when there is little additional variability among data sets and 2) to incorporate the variability among data sets into the overall estimate of uncertainty (standard errors or confidence limits). By combining potency estimates across data sets within studies and across studies, maximizes the use of the available information; almost all of the available dose-response data was used. Finally, this approach allows for a test of the dose-additivity assumption based on the similarly shaped dose-response curves because it generally forces an examination of each dose-response curve.

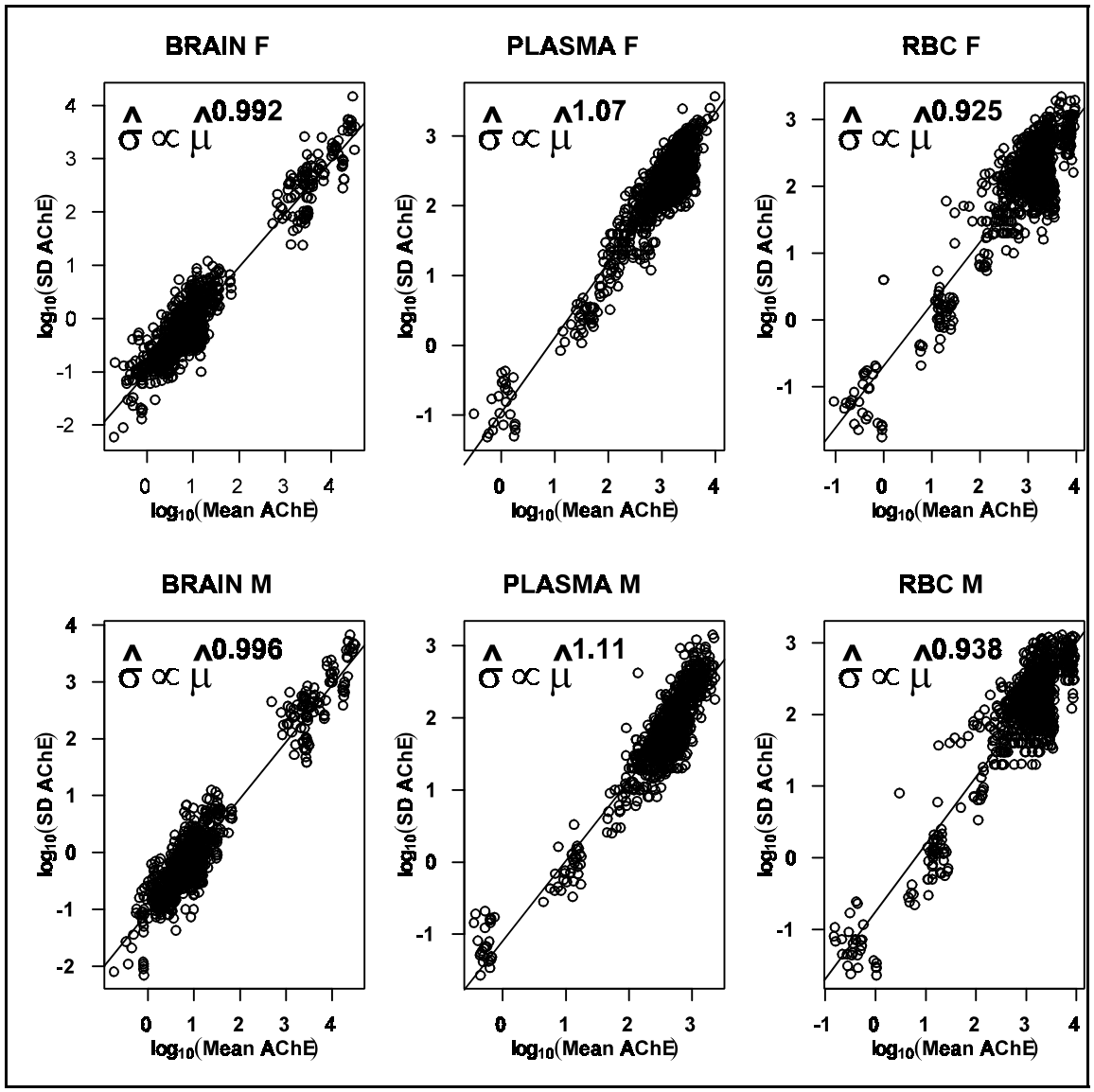

Figure III.B.1-1: Log_{10} of the standard deviation of cholinesterase activities plotted against the log_{10} of the corresponding means, for each compartment and sex. Each point is a single dose group. The plotted line is a regression line fitted to the data in each panel. The regression suggests the indicated relationship between estimated standard deviation and mean.

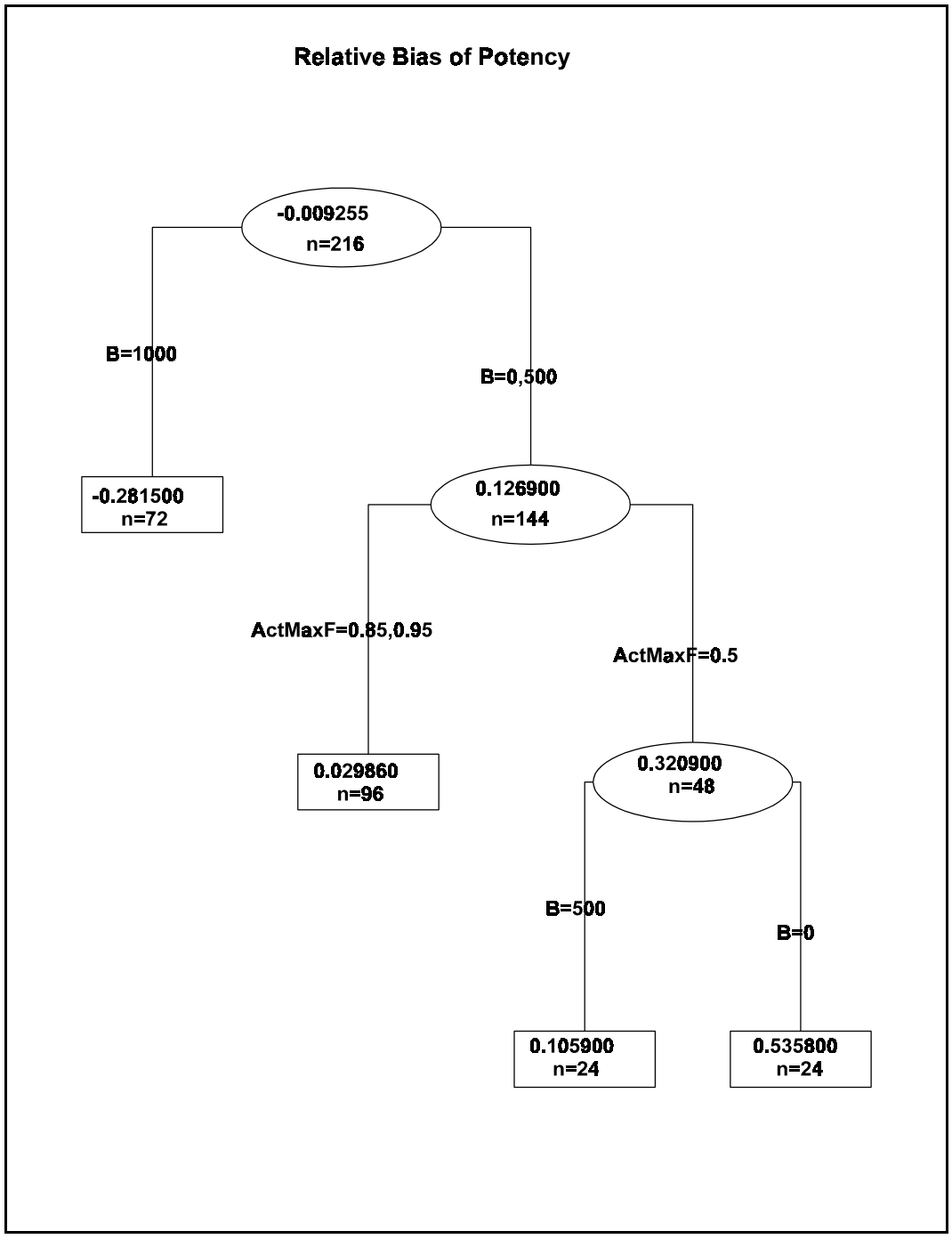

Figure III.B.1-2: Regression tree relating relative bias in potency estimates to model parameter values. The overall average relative bias is -0.009255 over all 216 combinations of conditions. If B < 1000 (that is B/A < 0.5), and the activity at the highest dose is nearly B, then the relative bias is only 0.0299.

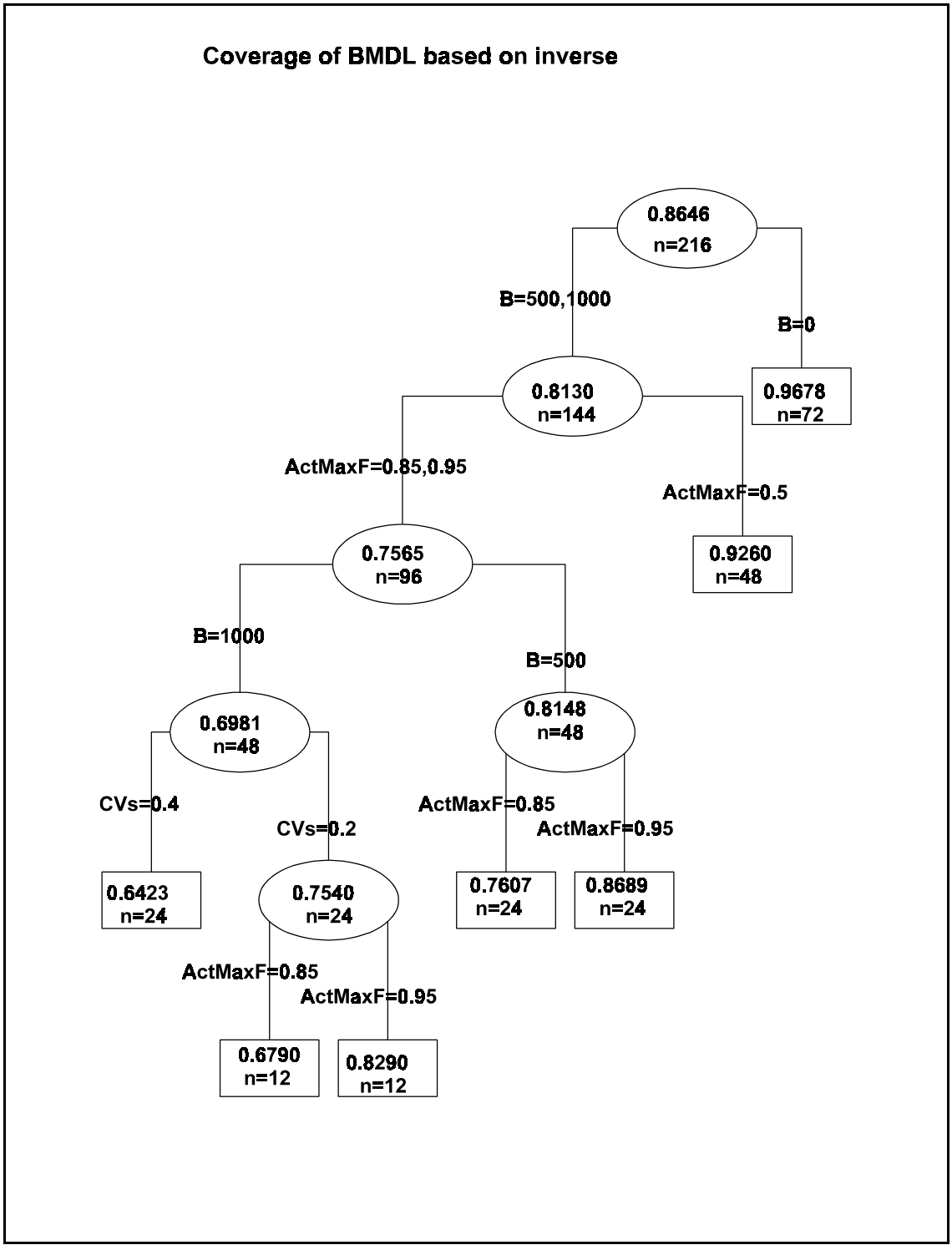

Figure III.B.1-3: Regression tree relating Coverage of the nominal 95% confidence limit for the BMD to the simulation conditions. The top number in each node is the average coverage at that node; n gives the number of observations that contribute to the average.

b. Follow-up Joint Analysis (December 2001)

i. Accounting for First-Pass Metabolism

When organophosphate pesticides (and many other chemicals) are administered orally, much of the absorbed chemical is carried to the liver by the portal circulation, where it may be metabolized. In the presence of saturable metabolism the dose-response curve would be expected to have a shallower slope at lower doses, and the slope would gradually increase as metabolism became saturated and more of the active chemical enters the general circulation. Although a detailed treatment of this process for each chemical is beyond the scope of this project, this basic idea was used to derive a two-parameter function of dose that relates administered dose to internal dose. The resulting function was combined with the basic exponential model giving a model that has a low dose shoulder while retaining the dose-response shape of the basic model for larger doses.

Consider the two-compartment pharmacokinetic model illustrated in Figure III.B.1-4:

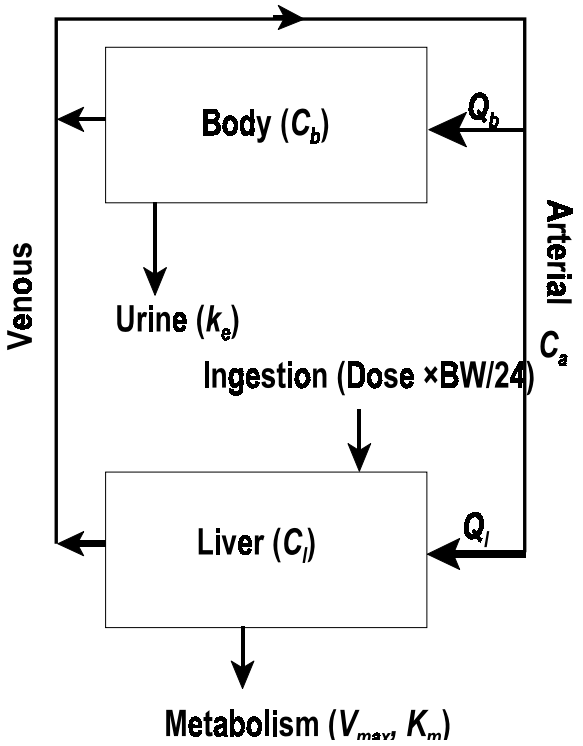

Figure III.B.1-4: Diagram for two-compartment PBPK model for the extension to the basic model.

In this simple model, all the ingested chemical is taken directly to the liver, where it is metabolized. The residual unmetabolized chemical is then distributed to the rest of the body through the circulation. Intake of chemical is continuous. In this case, two differential equations and one algebraic equation describe the concentration in the liver and the rest of the body:

$$V_b \frac{dC_b}{dt} = Q_b \times (C_a - C_b) - k_e C_b$$

$$V_l \frac{dC_l}{dt} = Q_l \times (C_a - C_l) + \frac{Dose \times BW}{24} - \frac{V_{\max} C_l}{K_m + C_l}$$

$$C_a = \frac{Q_b C_b + Q_l C_l}{Q_b + Q_l}$$

Here, C_x is the concentration in compartment x, where x is a for arterial blood, b for the body other than liver, and l for liver. The volume of and blood flow to compartment x are V_x and Q_x, where x is either b or l. V_{max} and K_m describe saturable metabolism of the chemical in the liver. The constant k_e is a first-order clearance term. Dose is expressed in milligrams per kilogram per day (hence the constant "24" to convert to hours), and body weight is expressed in kilograms. Thus, volumes in this parametrization are expressed in liters and concentrations in milligrams per liter.

At steady state, the derivatives are both 0: clearance just balances the dose rate. It can be shown (by solving the system of equations with derivatives set to zero) that the concentration in the body (C_b) at steady state is:

$$C_b = 0.5 * \frac{BW}{24 \times k_e} \left\{ \left(Dose - \frac{24 Q_l Q_b K_m k_e}{BW (Q_l k_e + Q_b k_e + Q_l Q_b)} - \frac{24 V_{\max}}{BW} \right) + \sqrt{\left(Dose - \frac{24 Q_l Q_b K_m k_e}{BW (Q_l k_e + Q_b k_e + Q_l Q_b)} - \frac{24 V_{\max}}{BW} \right)^2 + 4 Dose \frac{24 Q_l Q_b K_m k_e}{BW (Q_l k_e + Q_b k_e + Q_l Q_b)}} \right\}$$ **Eqn III.B.1-6**

Here, the odd constants 0.5 and 4 arise because the solution involves finding the roots of a quadratic polynomial, and 24 arises because dose rates are usually expressed in terms of "per day", while other coefficients in the model are "per hour".

Equation III.B.1-6 suggests using the function:

$$idose = 0.5 * \left\{ (Dose - S - D) + \sqrt{(Dose - S - D)^2 + 4 \times Dose \times S} \right\} \quad \textbf{Eqn III.B.1-7}$$

to describe the relationship between administered dose (*Dose*) and a scaled internal dose, where

$$S = \frac{24 Q_l Q_b K_m k_e}{BW\,(Q_l k_e + Q_b k_e + Q_l Q_b)},$$

and

$$D = \frac{24 V_{\max}}{BW}.$$

In this parameterization of the model, V_{max}, k_e, and total blood flow (= Q_b + Q_l) should be proportional to body weight, so both *S* and *D* are independent of body weight. This is a function of two parameters (*S* and *D*), and approaches the function *idose* = *Dose* – *D* for larger doses; the slope with respect to dose when *Dose* is close to 0 is *S/(S + D). D* quantifies the displacement of the relationship between *Dose* and *idose* from the identity relationship, and *S* controls the shape of the relationship at low doses. In the limit as $D \to 0$ or $S \to \infty$, Equation III.B.1-7 converges to *idose = Dose*.

In fact, it is reasonable to use Equation III.B.1-7 to approximate the relationship between internal dose and administered dose in the chronic dosing setting, even in the absence of a detailed pharmacokinetic justification. The general properties of the equation capture the expected effects of first-pass metabolic clearance of an active compound: a shallow shoulder of the curve at lower doses, with a slope that increases to a limit as the dose increases. As long as *S* and *D* are non-negative, varying these two parameters should result in a good approximation to virtually any low-dose deviation due to metabolic clearance, at least at the resolution available in bioassay dose-response data.

ii. The Nonlinear Mixed Effects Model

Mixed effects models in statistics are models for data in which some parameters vary among subsets of the data. For example, in this analysis, The background level (*lA*), scale factor (*lm*), and horizontal asymptote (*tB*) for the basic and expanded nonlinear models were presumed to vary among studies and among datasets nested within studies, whenever there was more than one study or dataset. The estimation problem for mixed effects models is to estimate both the fixed effects (parameters that do not vary), and the distribution parameters (for example, means and standard deviations) for the random parameters.

Typically (is in *nlme*), the random distribution is assumed to be normal, though methods exist for more general specifications, and the means of the random effects are estimated along with the fixed effects parameters.

The *R* package *nlme* estimates parameters for nonlinear mixed effects models using the approach described in Lindstrom and Bates (1990). Davidian and Giltinan (1995, pp 164 – 174) give a good description of this model, where they refer to it as being based on "conditional first-order linearization". This approach involves approximating the nonlinear function using a Taylor expansion before carrying out maximum likelihood estimation. The implementation in *nlme* allows the fixed and random effects to be expressed as linear models of other independent variables. In this analysis, for example, *lm* was allowed to differ between sexes by modeling *lm ~ sex - 1*, where *sex* is a categorical variable in the data set that takes the values *"F"* or *"M"*. The term "*- 1* " indicates that an intercept term should not be fit for this model, so there will be an estimate of *lm* for each sex. See the code in appendix 4 for "BRAINfits2.R" to see how this was used in more detail.