Benchmark Concentrations for Methylmercury Obtained from the Seychelles Child Development Study

Kenny S. Crump,[1] Cynthia Van Landingham,[1] Conrad Shamlaye,[2] Christopher Cox,[3] Phillip W. Davidson,[4] Gary J. Myers,[5] and Thomas W. Clarkson[6]

[1]The K.S. Crump Group, Inc., ICF Kaiser, Ruston, Louisiana, USA; [2]Ministry of Health, Victoria, Mahé, Republic of Seychelles; [3]Department of Biostatistics, [4]Strong Center for Developmental Disabilities, School of Medicine, University of Rochester, Rochester, New York, USA; [5]Department of Neurology, University of Rochester School of Medicine and Dentistry, Rochester, New York, USA; [6]Department of Environmental Medicine, University of Rochester Medical School, Rochester, New York, USA

Methylmercury is a neurotoxin at high exposures, and the developing fetus is particularly susceptible. Because exposure to methylmercury is primarily through fish, concern has been expressed that the consumption of fish by pregnant women could adversely affect their fetuses. The reference dose for methylmercury established by the U.S. Environmental Protection Agency was based on a benchmark analysis of data from a poisoning episode in Iraq in which mothers consumed seed grain treated with methylmercury during pregnancy. However, exposures in this study were short term and at much higher levels than those that result from fish consumption. In contrast, the Agency for Toxic Substances and Disease Registry (ATSDR) based its proposed minimal risk level on a no-observed-adverse-effect level (NOAEL) derived from neurologic testing of children in the Seychelles Islands, where fish is an important dietary staple. Because no adverse effects from mercury were seen in the Seychelles study, the ATSDR considered the mean exposure in the study to be a NOAEL. However, a mean exposure may not be a good indicator of a no-effect exposure level. To provide an alternative basis for deriving an appropriate human exposure level from the Seychelles study, we conducted a benchmark analysis on these data. Our analysis included responses from batteries of neurologic tests applied to children at 6, 19, 29, and 66 months of age. We also analyzed developmental milestones (age first walked and first talked). We explored a number of dose–response models, sets of covariates to include in the models, and definitions of background response. Our analysis also involved modeling responses expressed as both continuous and quantal data. The most reliable analyses were considered to be represented by 144 calculated lower statistical bounds on the benchmark dose (BMDLs; the lower statistical bound on maternal mercury hair level corresponding to an increase of 0.1 in the probability of an adverse response) derived from the modeling of continuous responses. The average value of the BMDL in these 144 analyses was 25 ppm mercury in maternal hair, with a range of 19 to 30 ppm. *Key words*: benchmark dose, child development, fetal exposure, fish, mercury, methylmercury, neurologic development, Seychelles. *Environ Health Perspect* 108:257–263 (2000). [Online 3 February 2000]
http://ehpnet1.niehs.nih.gov/docs/2000/108p257-263crump/abstract.html

Methylmercury is a neurotoxin that at high exposures can cause microcephaly, cerebral palsy, seizures, mental retardation, and death (*1,2*). Several poisoning episodes have confirmed that the fetal brain is particularly susceptible to methylmercury. The most severe effects were seen in Minamata and Niigata, Japan (*3*), where children were born with severe cerebral palsy in a population that consumed seafood contaminated with methylmercury from an industrial operation. Because exposure to methylmercury is primarily through the consumption of fish, concern has been raised regarding the health of children whose mothers consumed fish during pregnancy.

A reference dose (RfD) for methylmercury of 0.1 µg/kg/day was established by the U.S. Environmental Protection Agency (EPA) (*4*). This RfD was based on a benchmark analysis of data from a study of 81 Iraqi children whose mothers consumed seed grain treated with methylmercury as a fungicide (*5*). In this study an Iraqi mother's peak hair mercury concentration during pregnancy was associated with her child's neurodevelopment defined by the age at which the child first walked, first talked, and by a score derived from a neurologic examination. Although this study provided insight into exposure levels that can cause neurologic effects in the developing fetus, it has several features that may limit its usefulness in determining a dose response for chronic environmental exposure to methylmercury, such as might occur from eating fish. The ages at which the Iraqi children first walked or first talked were based on family members' memories obtained at a mean age of 19 months after delivery, and were not known precisely. Even birth dates were not recorded precisely but were established in relation to religious holidays, seasons, etc. There was limited information available on the child's environment, which could have influenced his or her neurologic development. The measures of effect in the study (late walking, late talking, and neurologic score) were relatively nonspecific to neurologic deficit and may not have been sensitive to methylmercury. Perhaps most important, the exposures in the Iraqi study occurred at a high level for a limited period of time, and consequently were not typical of lower level and often chronic exposures associated with fish consumption.

In contrast to the EPA's use of the Iraqi study, the Agency for Toxic Substances and Disease Registry (ATSDR) (*6*) developed its proposed minimum risk level (MRL) of 0.3 µg/kg/day for methylmercury from the Seychelles Child Development data (*7,8*). The Seychelles study examined longitudinally the neurologic and psychologic responses of 708 children whose mothers consumed a relatively large amount of fish during pregnancy (*7,8*). The children were given a range of sensitive tests of neurologic and psychologic development. Information was available on a number of health, social, and demographic variables that could have been associated with test performance. Most important, exposures in this study were through fish, which is the route of concern in human populations. Because there was no association between methylmercury exposure and neurologic or psychologic outcomes,

Address correspondence to K. Crump, The K.S. Crump Group, Inc., ICF Kaiser, 602 East Georgia Avenue, Ruston, LA 71270 USA. Telephone: (318) 242-5019. Fax: (318) 255-4960. E-mail: kcrump@iamerica.net

Support for this work was provided by the Electric Power Research Institute, Palo Alto, California.

Received 18 May 1999; accepted 2 September 1999.

ATSDR investigators considered the mean maternal hair concentration of 15.3 ppm from the Seychelles group with the highest exposure to be a no-observed-adverse-effect level (NOAEL), and used that NOAEL as its basis for deriving the MRL. In an alternative derivation of its MRL, ATSDR investigators considered the mean maternal hair concentration of 6.8 ppm from the entire Seychelles cohort at 66 months to be a NOAEL.

The NOAEL approach applied by the ATSDR is the traditional method for deriving an exposure judged to be without appreciable risk of noncancer health effects to humans (e.g., an RfD or an MRL). This approach is most appropriate for experimental studies in which animals are placed in a few discrete dose groups. Determination of a NOAEL is more problematic in epidemiologic studies such as the Iraqi or Seychelles studies, in which exposures do not assume a few discrete values. When an effect is seen in an epidemiologic study, a mean exposure has often been assumed to be a lowest-observed-adverse-effect level (LOAEL). However, the analysis may have demonstrated only that an effect occurred somewhere within the range of exposures in the study—a range that might span several orders of magnitude. Similarly, when an effect is not seen in an epidemiologic study, a mean exposure is sometimes assumed to be a NOAEL, as was the case in the use of the Seychelles study by ATSDR investigators. However, the mean exposure may not be a good indicator of the no-effect exposure. For example, adding a large number of individuals to the cohort with very low exposures could dramatically affect the mean exposure, but would add little information on the appropriate value for a NOAEL.

The benchmark approach applied to the Iraqi data by EPA investigators provides an alternative to the NOAEL in determining an RfD (*9–11*). The benchmark dose (BMD) is the dose that corresponds to a specified level of additional response called the benchmark response (BMR). The BMD is calculated by fitting a mathematical dose–response model to the data. A lower statistical confidence bound on the BMD (BMDL) replaces the NOAEL in the calculation of an RfD. In addition to methylmercury, as documented in the EPA Integrated Risk Information Service (IRIS database) (*12*), the EPA has used this approach to determine RfDs or reference air concentrations for a number of other chemicals.

The BMD method has several advantages over the NOAEL method, including making better use of dose–response information and reflecting sample size more appropriately (*9,11*). When applied to epidemiologic data, the BMD allows for consideration of the dose response over the entire exposure range rather than assuming that the mean exposure is either a NOAEL or LOAEL.

Another advantage of the BMD method is that a BMDL can be calculated from negative data (data in which no statistically significant dose-related trend is present). With negative data, the point estimate of the dose–response trend may be positive (in the direction of an adverse effect), zero, or negative (in the direction of a beneficial effect), although any deviation from zero may reflect only random variation and not a real effect of exposure. If the point estimate of the trend is zero or negative, the point estimate of the BMD will be infinite (undefined), whereas if the point estimate of the trend is positive, the point estimate of the BMD will be finite. A statistical upper bound on the BMD obtained from negative data will be infinite; otherwise the data would, by definition, not be negative. However, a BMDL derived from negative data will be a finite number as long as the data do not demonstrate a statistically significant negative trend. When data are negative, it is possible that there is no effect of treatment, in which case a BMDL reflects only the statistical constraints imposed by the experimental design. However, a BMDL represents a conservative (in the health-protective sense) value. Even though the data were negative, exposure could have caused a small undetected increase in an adverse health effect. The BMDL is a statistical bound that reflects the potential size of such an increase.

We present benchmark analyses of the Seychelles data similar to those applied to the Iraqi data by the EPA in the development of the methylmercury RfD. These analyses may provide a sounder scientific basis for developing an RfD or MRL from the Seychelles data than a NOAEL estimated from this study. In addition to the Seychelles study, neurologic studies of children from high fish-consuming populations have also been conducted in New Zealand (*13,14*) and the Faroe Islands (*15*). These studies should also be considered when developing an exposure advisory for methylmercury. A benchmark analysis similar to ours from the Seychelles data has been developed from the New Zealand data (*14*).

The Seychelles Study

The Seychelles Child Development Study (*7,8,16,17*) began in 1989 in the Republic of Seychelles, an archipelago in the Indian Ocean, to measure developmental outcomes in children whose mothers consumed fish containing methylmercury during their pregnancy. More than 700 mother/infant pairs were enrolled in the main study during pregnancy. The children's prenatal mercury exposure was assessed by measuring the concentration of mercury in a segment of maternal hair representative of the hair formed during pregnancy. At 6.5 months of age, the children were given a neurologic examination, tests of visual recognition memory (*18*) and visual attention (*19*), and the Denver Developmental Screening Test - Revised (*20*). The Bayley Scales of Infant Development (*21*) were given at 19 months and again at 29 months. The ages at which a child first walked and first talked were also recorded. At 66 months of age the children were given a battery of tests that measured cognitive ability [McCarthy Scales of Children's Abilities (*22*)], language ability [Preschool Language Scale (*23*)], reading and arithmetic achievement [Woodcock-Johnson Achievement Tests (*24*)], visual–spatial adeptness [Bender-Gestalt Test (*25*)], and social behavior [Child Behavior Checklist (*26*)]. Covariables likely to be associated with a child's development were also assessed, including maternal intelligence [Raven Standard Progressive Matrices (*27*)], home environment [HOME (*28*)], parental education, and others shown in Table 1.

Mercury hair concentrations in maternal hair ranged from 0.5 to 26.7 ppm, and averaged 6.8 ppm. By contrast, hair mercury concentrations in the U.S. are estimated to average 0.5 ppm (*4*). No adverse outcomes were associated with mercury exposure at any age in the Seychelles study.

Benchmark Methodology

Most of the measured end points in the Seychelles study were recorded as continuous responses, and several dose–response models were applied to these continuous end points. In the *k*-power model (*10*), the mean test score (or a transformation thereof) was assumed to be of the form

$$\mu(d) = \pm (\beta d)^K + \beta_0 + \beta_1 C_1 + \ldots + \beta_m C_m, \quad [1]$$

where d is the average maternal hair mercury level during pregnancy in milligrams per kilogram, $C_1, \ldots, C_m$ are covariates that may also be correlated with test scores, and $\beta \geq 0$, $\beta_0, \ldots, \beta_m$, and $K \geq 1$ are parameters to be estimated. The negative sign was assumed whenever smaller responses were considered more adverse. We also assumed that the test scores were normally distributed (possibly after transformation) with a mean, $\mu(d)$, given by Equation 1 and a SD, σ, that was independent of all covariates, including the hair mercury concentration. Inclusion of the parameter, K, allowed a nonlinear dose response. If K is fixed at $K = 1$, the *k*-power model is identical to that used in ordinary multiple linear regression.

If test scores below a predetermined value, x_0, are considered abnormal, the probability

of an abnormal test score from a child with a mercury hair concentration of d mg/kg is

$$P(d) = N\{[x_0 - \mu(d)]/\sigma\}, \quad [2]$$

where N is the standard normal cumulative distribution function. The probability of an abnormal test score in an unexposed child is

$$p_0 \equiv P(0) = N\{[x_0 - \mu(0)]/\sigma\}. \quad [3]$$

Similar formulas apply when larger scores are more adverse (*10*). The BMD was defined as the maternal mercury hair concentration that caused the probability of an abnormal test score to increase by an amount BMR; i.e., as the hair concentration that satisfied

$$P(\text{BMD}) - p_0 = \text{BMR}. \quad [4]$$

Although it is not explicit in the notation, whenever a single cutoff, x_0, is used to define an abnormal response, both p_0 and BMD will be dependent on the covariates.

Alternatively, instead of specifying the cutoff, x_0, for an abnormal response, it is possible to specify the percentage of unexposed individuals who have an abnormal response, i.e., specify p_0. If p_0 is specified under the k-power model, the definition of the BMD (*4*) is equivalent to defining the BMD as the exposure that corresponds to a given change (BMRC) in the mean response normalized by the SD; i.e., as the exposure that satisfies

$$[\mu(\text{BMD}) - m(0)]/\sigma = \text{BMRC} \quad [5]$$

(*10*). For given values for p_0 and BMR used in Equation 4 there is a corresponding value for BMRC in Equation 5 so that Equations 4 and 5 are equivalent. For example, for p_0 = 0.05 and BMR = 0.1, the corresponding value for BMRC = 0.61. Unlike when x_0 is specified, when p_0 is specified, the BMD does not depend on the covariates (*10*). This is obvious from the equivalent definition of the BMD given by Equation 5.

In addition to the k-power model, two other dose–response models were applied to continuous data from the Seychelles study. As previously noted, if the underlying data are normally distributed with a common SD, specification of a dose–response model, $\mu(d)$, for the mean response and specification of a dividing point, x_0, between normal and abnormal (or, alternatively, of a proportion, p_0, of unexposed subjects whose responses are considered abnormal) determines the probability, $P(d)$, of an adverse response as a function of exposure. Alternatively, we can specify x_0 or p_0 and a functional form for the probability, $P(d)$, of an adverse response; this determines the mean response, $\mu(d)$, where $\mu(0)$ is a linear function of the covariates, as in the k-power model. This approach allows us to define a model for continuous data that corresponds to a predetermined functional form for the probability of an adverse response, and consequently permits the same mathematical dose–response model to be applied to quantal and continuous data (*10*).

Two such dose–response models for the mean response were applied to the Seychelles data. One model was determined by the Weibull model for quantal data,

$$P(d) = p_0 + (1 - p_0)\{1 - \exp[-(\beta d)^k]\}, \quad [6]$$

and one determined by the logistic model,

$$P(d) = p_0 + (1 - p_0)\{1 - 1/[1 + (\beta d)^K]\}, \quad [7]$$

where $\beta \geq 0$ and $K \geq 1$ are estimated parameters that have roughly the same interpretation as the corresponding parameters in the k-power model. As in the k-power model, with these models, the BMD will be a function of the covariates when x_0 is specified, but not when p_0 is specified.

Each of the three models for continuous data (k-power, Weibull, and logistic) was applied to the continuous responses from Seychelles data both with p_0 specified and with x_0 specified. In addition, a linear model was applied. The linear model was defined by fixing $K = 1$ in the k-power model. In analyses in which p_0 was specified, it was fixed at 0.05, as suggested by the convention of considering 95% of clinical responses in healthy individuals to define the normal range. In analyses in which x_0 was specified, it was fixed at 2 SDs (in the adverse direction) from the overall mean response.

For a few of the tests a child's response was recorded as a quantal response (e.g., abnormal/normal), and the results from these tests were modeled using the Weibull (Equation 6) dose–response model for quantal data. In addition, each of the continuous responses was converted into a quantal response by considering a response abnormal if it was more than 2 SDs away (in the adverse direction) from the mean response of the entire cohort, and then analyzed using the Weibull model. In these analyses the BMD was defined in the same way as in the analyses of the continuous responses, i.e., as the mercury hair concentration that corresponded to an increase in 0.1 in the probability of an abnormal response [Equation 4 with BMR = 0.1].

We conducted each analysis of a continuous response without covariates. In addition, analyses with p_0 specified were also conducted using both an expanded set and a reduced set of covariates (Table 1). These sets of covariates have essentially the same definitions as those in the original analyses (*7,8,16,17*). In all models for continuous responses with p_0 specified, the BMD, as well as corresponding point estimate and confidence limit, does not depend on the covariates. Covariates were not included in analyses of quantal responses or in analyses of continuous responses in which x_0 was

Table 1. Definitions of reduced (red) and expanded (exp) sets of covariates used in the benchmark calculations.

	6 months		19 and 29 months		First walked or talked		66 months	
Covariate	Red	Exp	Red	Exp	Red	Exp	Red	Exp
Sex	X	X	X	X	X	X	X	X
Birth weight[a]	X	X	X	X	X	X	X	X
Birth order[a]	X	X	X	X	–	X	X	X
Breast fed[b]	X	X	X	X	–	X	X	X
Child's medical history[c]	X	X	X	X	X	X	X	X
Maternal age[a]	X	X	X	X	X	X	X	X
Maternal smoking in pregnancy[d]	–	X	–	X	–	X	–	X
Maternal alcohol in pregnancy[d]	–	X	–	X	–	X	–	X
Maternal medical history[e]	X	X	X	X	–	X	X	X
Language spoken in home (home)[f]	X	X	X	X	X	X	X	X
Score from home visit[g]	–	–	X	X	X	X	X	X
Raven group (caregiver IQ)[h]	–	X	–	X	–	X	–	X
Maternal education, level[i]	X	X	X	X	X	X	–	–
Paternal education, level[i]	X	X	X	X	–	X	–	–
Family income[j]	–	X	–	X	–	X	–	–
Gestation age[a]	–	X	–	X	–	X	–	–
Hollingshead socioeconomic scale[k]	–	–	–	–	–	–	X	X
Hearing[l]	–	–	–	–	–	–	X	X
Child's mercury level[a]	–	–	–	–	–	–	X	X

IQ, intelligence quotient.
[a]Modeled as a continuous variable. [b]0–1 month, 1–6 months, or > 6 months. [c]Negative, diagnosed intrauterine growth retardation or head circumference < or > 2 SD from normal. [d]No or yes. [e]Reported hypertension requiring medication, negative. [f]Creole, other language. [g]≤ 31, > 31–35, or > 35. [h]Low < 16, normal 16–28, or high > 28. [i]primary education, secondary through trade school, or polytechnic through university. [j]500–1,500 [Seychelles rupee (SR)]; 1,501–3,000 SR; 3,001–4,000 SR; or > 4,000 SR. [k]Unskilled ≤ 19, semiskilled > 19–29, skilled > 29–39, or minor/major business/professional > 39. [l]≤ 25 db, > 25–35 db, or > 35 db.

specified because, in both of these cases, the BMD would be different for different values of the covariates; therefore, to calculate a BMDL using covariates, values would have to be specified for each of the covariates. As in the original analyses, certain responses were transformed to make them conform more closely to a normal distribution, and outliers were eliminated. A logarithmic transformation was applied to outcomes of age first walked and first talked, and a negative reciprocal transformation was applied to the data on the Fagan attention outcome (*18*) at 6 months. The data from the psychomotor index were highly skewed at both 19 and 29 months, and were only analyzed in quantal form. An outlier was defined as a response for which the absolute value of the residual was greater than 3 times the SD of the residuals.

Parameter estimates were obtained using the maximum likelihood method, and statistical confidence bounds (e.g., the BMDL) were computed by the profile likelihood method (*29*). The BMDL was defined conventionally as the 95% statistical lower confidence bound on the BMD. In the implementation of the profile likelihood method, a model was first reparameterized so that the BMD was included explicitly as a parameter in place of β. Next, values (maximum likelihood estimates) of the parameters were obtained so that the log-likelihood, L, of the data obtained its largest possible value, L_0. The maximum likelihood estimate of the BMD could theoretically be infinite. The BMDL was then determined as the smallest value of the BMD that satisfied the equation, $\{2[L_0 - L(\text{BMD})]\}^{1/2} = 1.645$, where 1.645 is the 95th percentile of the standard normal distribution, and L(BMD) is the function of the BMD (only) obtained by fixing the value of the BMD and maximizing the log-likelihood with respect to the remaining parameters. These calculations were made using computer programs specifically designed for performing benchmark calculations (*30*).

Results

Table 1 provides specific BMDLs obtained from each neurologic test using the Weibull model (applied to both continuous and quantal responses) and the *k*-power model. Table 2 provides averages and ranges of BMDLs over all neurologic tests for each type of benchmark analysis applied.

Figure 1 illustrates the calculation of the BMDL for five continuous responses, based on the *k*-power model, with p_0 specified and using the full set of covariates. In this case (using $p_0 = 0.05$ and BMR = 0.1), an equivalent definition of the BMD is the hair concentration that causes the mean response to differ from the mean response of unexposed individuals, relative to σ, by BMRC = 0.61 [Equation 5]. Each plotted point in Figure 1 represents a residual (measured response minus the mean response predicted by the model) with the mercury effect and the constant term added back in [i.e., (residual) + $\beta_0 \pm (\beta d)^K$]. The solid line represents the estimated mean response as a function of hair mercury, $\beta_0 \pm (\beta d)^K$. The dotted curve represents the statistical bounding mean dose response associated with the BMDL (i.e., $\beta_0 \pm (\beta' d)^{K'}$, where β', and K' are the values of β and K associated with the calculation of the BMDL). The dotted line may be used to approximate BMDLs associated with other values of p_0 and BMR by using the corresponding value for BMRC. However, this approach provides only an approximation because β', and K' may be slightly different for different values of p_0 and BMR. Values of BMRC that correspond to some specific values for p_0 and BMR are shown in Table 3. Equivalent BMRCs that correspond to other values of p_0 and BMR may be obtained using the formula BMRC = $N^{-1}(1 - p_0) - N^{-1}(1 - p_0 - \text{BMR})$, where N^{-1} is the inverse of the standard normal distribution function (*10*).

Discussion

The point estimates of the BMDs (data not shown) were in many instances theoretically infinite because many of the dose–response trends, although nonsignificant, were in the direction of a beneficial effect of mercury. However, the BMDLs were all finite, and their values reflect the potential magnitude of any small effect of mercury that went undetected in this study.

BMDLs obtained from continuous data using the Weibull, logistic, and *k*-power models were similar. Application of the linear model resulted in larger BMDLs (data not shown) that were outside the range of maternal hair levels observed in the study. The inclusion of covariates had little effect on the BMDLs. Likewise, BMDLs were similar whether based on specification of the background response, p_0, or of x_0, the magnitude of a response considered abnormal. However, BMDLs based on quantal data were consistently smaller than those based on the corresponding continuous data. For example, the BMDLs computed from 12 quantal responses using the Weibull distribution (Table 4) had a mean of 22.0 ppm and a range of 19.4–23.7 ppm, whereas the 12 BMDLs computed in like manner (Weibull distribution, no covariates) from the corresponding continuous data had a mean of 25.0 ppm and a range of 23.1–27.2 ppm. One possible reason for this difference is that collapsing a continuous data point into a quantal (yes/no) response results in a loss of information, which can cause confidence intervals on the BMD to become wider (i.e., cause the BMDL to become smaller). For this reason, we believe that a BMDL derived from a continuous response is preferred over one derived from a quantal end point obtained by collapsing the corresponding continuous response, provided the distributional

Table 2. Some BMDLs (parts per million mercury in maternal hair) obtained using the Weibull and *k*-power models and listed by covariate.

	Weibull				*k*-Power		
	p_0[a]		x_0[b]	Quantal	p_0[a]		x_0[b]
Model	None	Exp	None	None	None	Exp	None
6 months							
Deep tendon reflexes	–	–	–	22.8	–	–	–
Limb tone	–	–	–	20.9	–	–	–
Overall neurologic	–	–	–	15.8	–	–	–
Fagan visual recognition memory	26.0	26.0	27.4	19.7	26.0	26.0	26.9
Fagan attention	25.7	25.9	27.0	23.7	25.5	25.6	26.4
19 months							
Mental development index	23.7	23.4	26.0	22.6	24.3	24.1	25.6
Psychomotor index	–	–	–	22.3	–	–	–
29 months							
Mental development index	24.1	24.4	25.7	21.9	24.0	24.2	24.8
Psychomotor index	–	–	–	22.5	–	–	–
66 months							
Bender Gestalt errors	26.9	26.7	28.5	22.7	26.7	26.7	27.5
Child Behavior Checklist total	27.2	27.2	29.0	19.4	20.0	26.9	27.8
McCarthy general cognitive index	24.4	24.2	26.5	22.7	24.7	24.6	25.9
Preschool Language total score	25.2	25.1	26.8	22.7	24.7	24.7	25.5
Woodstock-Johnson							
Applied problems	23.1	23.5	25.3	22.7	23.9	24.3	25.5
Letter word recognition	23.7	23.7	25.3	22.7	23.8	23.9	24.7
Developmental milestones							
Age first walked unassisted	24.9	24.0	25.9	22.7	24.4	23.2	26.8
Age first talked	24.6	23.5	25.9	20.3	25.0	24.1	25.9

Exp, expanded.
[a]Abnormal defined as a response > 2 SDs in adverse direction from mean response of entire cohort. [b]Abnormal defined so that 5% of responses are abnormal ($p_0 = 0.05$).

assumption (e.g., normality) required for a continuous analysis is satisfied. The 144 BMDLs derived from continuous data using the Weibull, logistic, or *k*-power models averaged 25.0 ppm with a range of 19.4–30.3 ppm (Table 4).

Several choices were necessary to calculate our BMDLs, including how the background response was specified (e.g., the values of p_0 or x_0 selected), the value selected for the benchmark risk (BMR), and the dose–response model used. These decisions are generic in the sense that they are largely unrelated to the type of exposure, health response, or quality of data. The Weibull dose–response model was used by the EPA to determine the current RfD (*4*), and was also used in a number of benchmark analyses listed in the IRIS database (*12*). Our choice of BMR = 0.1 was used by the EPA to derive its RfD for mercury, and was used in all of the benchmark analyses presently in the EPA IRIS database (*12*). Thus, a precedent appears to have developed regarding the choice for the BMR. Our choice of $p_0 = 0.05$ (equivalent to assuming that 5% of persons from a healthy unexposed population are abnormal) is consistent with the definition of the normal range in clinical data; therefore, we believe this is a reasonable choice, although perhaps at the upper end of reasonable. The first paper that discussed this approach proposed using $p_0 = 0.001$ (*31*). Generally speaking, smaller values of p_0 result in larger BMDLs.

Except for the linear model, which produced larger BMDLs, the dose–response models applied to continuous end points (Weibull, logistic, and *k*-power) all produced comparable BMDLs. This was not surprising because these dose responses are fairly similar. However, it is possible for different dose–response models to yield radically different BMDLs. This is a problem, particularly when the dose response is permitted to rise very steeply in the low-dose range (i.e., to have unboundedly large slopes at increasingly small doses). Not only can this type of dose–response model yield ridiculously low BMDLs, the dose–response represented by such a model is biologically implausible (*10*). We avoided this problem in our models by restricting the parameter, *K*, to no less than 1.0. (Any value of $K < 1$ produces an infinite slope at zero dose.)

Kjellström et al. (*13*) administered scholastic and psychologic tests to 237 6- and 7-year-old children in New Zealand and correlated the results with the mercury concentration in their mothers' hair during

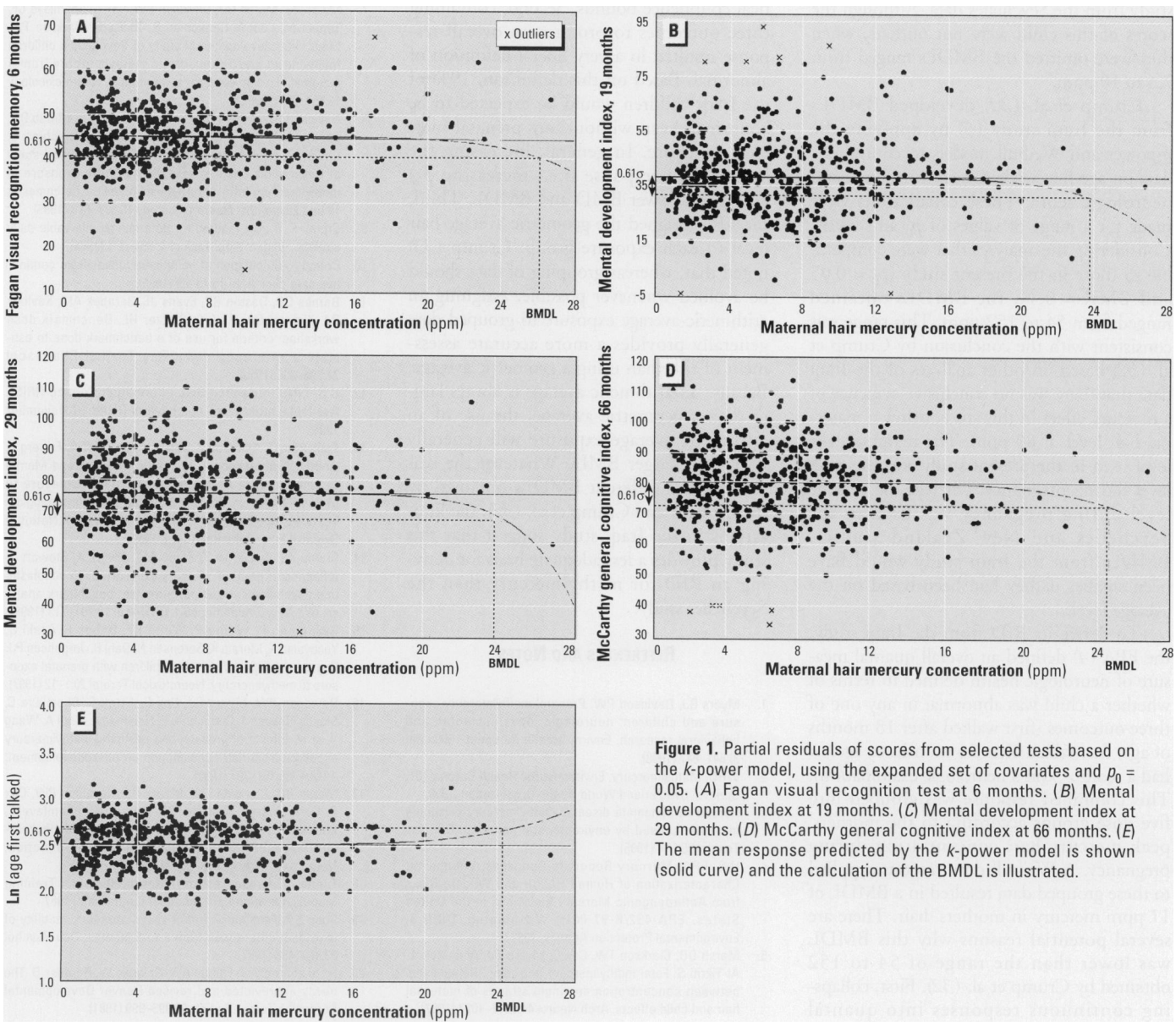


Figure 1. Partial residuals of scores from selected tests based on the *k*-power model, using the expanded set of covariates and $p_0 = 0.05$. (*A*) Fagan visual recognition test at 6 months. (*B*) Mental development index at 19 months. (*C*) Mental development index at 29 months. (*D*) McCarthy general cognitive index at 66 months. (*E*) The mean response predicted by the *k*-power model is shown (solid curve) and the calculation of the BMDL is illustrated.

Table 3. Values of BMRC corresponding to certain values for p_0 and BMR.

p_0	BMR	BMRC
0.01	0.05	0.77
0.01	0.1	1.1
0.05	0.05	0.36
0.05	0.1	0.61

pregnancy. Crump et al. (*14*) performed benchmark calculations on five end points from this study by applying the *k*-power model in the same manner as in the present paper. The results of this analysis were highly dependent on the test scores of one child whose mother had a hair mercury level of 86 ppm during pregnancy, which was more than 4 times the hair mercury level of any other mother. When this child's scores were included in the analysis, BMDLs ranged from 17 to 24 ppm, which is similar to the range of BMDLs determined in the present study from the Seychelles data. Although the scores of this child were not outliers, when they were omitted the BMDLs ranged from 7.4 to 10 ppm.

Crump et al. (*32*) developed BMDLs from the Iraqi study (*5*) by applying the *k*-power and Weibull models to continuous data on age first walked, age first talked, and neurologic score. These calculations were made for a range of values of p_0 and BMR. Considering the analyses that were comparable to those in the present study (p_0 = 0.05 and BMR = 0.1), the BMDLs obtained ranged from 54 to 152 ppm. This range was consistent with the conclusion by Crump et al. (*32*) based on other analyses of the Iraqi data that there was no conclusive evidence of a mercury effect in this study below a maternal hair level of 80 ppm. The maternal hair level used in the present study was the peak level during pregnancy, whereas the average level during pregnancy was used in the Seychelles and New Zealand studies. BMDLs from the Iraqi study would have been smaller if they had been based on the average level.

To derive its RfD from the Iraqi study, the EPA (*4*) defined an overall quantal measure of neurologic health defined in terms of whether a child was abnormal in any one of three outcomes (first walked after 18 months of age, first talked after 24 months of age, or had a score > 3 on a neurologic examination). This combined response was grouped into five dose groups according to the mothers' peak mercury hair concentration during pregnancy. A Weibull dose response applied to these grouped data resulted in a BMDL of 11 ppm mercury in mothers' hair. There are several potential reasons why this BMDL was lower than the range of 54 to 152 obtained by Crump et al. (*32*). First, collapsing continuous responses into quantal responses results in a loss of information and consequently generally produces wider statistical confidence bounds. Second, combining three outcomes to produce an overall response resulted in a very liberal definition of abnormal. Based on this definition, 19% of the Iraqi children would be expected to be abnormal, even without any prenatal mercury exposure. In general, increasing the background response (i.e., increasing p_0) results in a lower BMD and BMDL. Third, the EPA assigned the geometric average hair level to each exposure group. Crump (*33*) noted that, whereas grouping of data should be avoided whenever possible, assigning an arithmetic average exposure to grouped data generally provides a more accurate assessment of risk than using a geometric average. Because an arithmetic average is always larger than a geometric average, the use of an arithmetic average exposure will generally result in a larger BMD. Whatever the reasons for the different BMDLs obtained by the EPA (*4*) and Crump et al. (*32*), the limitations in the Iraqi study suggest that this study provides a less adequate basis for deriving an RfD for methylmercury than the Seychelles study.

Table 4. Summary of BMDLs (parts per million mercury in maternal hair) by BMD analysis method.

Type of data	Definition of abnormal	Covariates	Model	No.	Average	Range
Continuous	x_0[a]	None	Weibull	12	26.6	25.3–29.0
			Logistic	12	26.7	25.3–29.1
			k-Power	12	26.1	24.7–27.8
	p_0[b]	None	Weibull	12	25.0	23.1–27.2
			Logistic	12	25.0	23.0–27.3
			k-Power	12	24.4	20.0–26.7
		Reduced	Weibull	12	25.4	23.5–30.1
			Logistic	12	25.4	23.3–30.3
			k-Power	12	24.5	19.4–26.8
		Expanded	Weibull	12	24.8	23.4–27.2
			Logistic	12	24.7	23.2–27.2
			k-Power	12	24.9	23.2–26.9
Quantal		None	Weibull	17	21.6	15.8–23.7

[a]Abnormal defined as a response > 2 SDs in adverse direction from mean response of entire cohort. [b]Abnormal defined so that 5% of the responses are abnormal (p_0 = 0.05).

References and Notes

1. Myers GJ, Davidson PW. Prenatal methylmercury exposure and children: neurologic, developmental, and behavioral research. Environ Health Perspect 106(suppl 3):841–847 (1998).
2. WHO. Methylmercury. Environmental Health Criteria 101. Geneva, Switzerland:World Health Organization, 1990.
3. Harada H. Minamata disease: methylmercury poisoning in Japan caused by environmental pollution. Crit Rev Toxicol 25:1–4 (1995).
4. U.S. EPA. Mercury Report to Congress, Volume VI: Characterization of Human Health and Wildlife Risks from Anthropogenic Mercury Emissions in the United States. EPA-452/R-97-001f. Washington, DC:U.S. Environmental Protection Agency, 1997.
5. Marsh DO, Clarkson TW, Cox C, Myers GJ, Amin-Zaki L, Al-Tikriti S. Fetal methylmercury poisoning. Relationship between concentration iin single strands of maternal hair and child effects. Arch Neurol 44:1017–1022 (1987).
6. ATSDR. Toxicological Profile for Mercury (Update). Atlanta, GA:Agency for Toxic Substances and Disease Registry, 1999.
7. Myers GJ, Marsh DO, Davidson PW, Cox C, Shamlaye CF, Tanner M, Choi A, Cernichiari E, Choisy O, Clarkson TW. Main neurodevelopmental study of Seychellois children following *in utero* exposure to methylmercury from a maternal fish diet: outcome at six months. Neurotoxicology 16:653–664 (1995).
8. Davidson PW, Myers GJ, Cox C, Shamlaye CF, Marsh DO, Tanner MA, Berlin M, Sloane-Reeves J, Cernichiari E, Choisy O, et al. Longitudinal neurodevelopmental study of Seychellois children following *in utero* exposure to methylmercury from maternal fish ingestion: outcomes at 19 and 29 months. Neurotoxicology 16:677–688 (1995).
9. Crump K. A new method for determining allowable daily intakes. Fundam Appl Toxicol 4: 854–871 (1984).
10. Crump K. Calculation of benchmark doses from continuous data. Risk Anal 15:79–89 (1995).
11. Barnes DG, Daston GP, Evans JS, Jarabek AM, Kavlock RJ, Kimmel CA, Park C, Spitzer HL. Benchmark dose workshop: criteria for use of a benchmark dose to estimate a reference dose. Reg Toxicol Pharmacol 21:296–306 (1995).
12. U.S. EPA. Integrated Risk Information Service (IRIS). Available: http://www.epa.gov/iris [cited 16 November 1998].
13. Kjellström T, Kennedy P, Wallis S, Stewart A, Friberg L, Lind B, Wutherspoon T, Mantell C. Physical and Mental Development of Children with Prenatal Exposure to Mercury from Fish. Stage 2: Interviews and Psychological Tests at Age 6. Report 3642. Solna, Sweden:National Swedish Environmental Board, 1989.
14. Crump KS, Kjellström T, Shipp AM, Silvers A, Stewart A. Influence of prenatal mercury exposure upon scholastic and psychological test performance: benchmark analysis of a New Zealand cohort. Risk Anal 18:701–713 (1998).
15. Grandjean P, Weihe P, White RF, Debes F, Araki S, Yokoyama K, Murata K, Sørensen N, Dahl R, Jørgensen PJ. Cognitive deficit in 7-year-old children with prenatal exposure to methylmercury. Neurotoxicol Teratol 20:1–12 (1997).
16. Davidson PW, Myers GJ, Cox C, Axtell C, Shamlaye C, Sloane-Reeves J, Cernichiari E, Needham L, Choi A, Wang Y, et al. Effects of prenatal and postnatal methylmercury exposure from fish consumption on neurodevelopment. JAMA 280:701–707 (1998).
17. Marsh DO, Clarkson TW, Myers GJ, Davidson PW, Cox C, Cernichiari E, Tanner MA, Lednar W, Shamlaye C, Choisy O, et al. The Seychelles study of fetal methylmercury exposure and child development: introduction. Neurotoxicology 16:583–596 (1995).
18. Fagan JF. Fagan Test of Infant Intelligence: Training Manual. Cleveland, OH:Infantest Corporation, 1987.
19. Rose SA, Feldman J. Infant visual attention: stability of individual differences from 6 to 8 months. Dev Psychol 23:490–498 (1987).
20. Frankenburg WK, Fandal AW, Sciarillo W, Burgess D. The newly abbreviated and revised Denver Developmental Screening Test. J Pediatr 99:995–999 (1981).

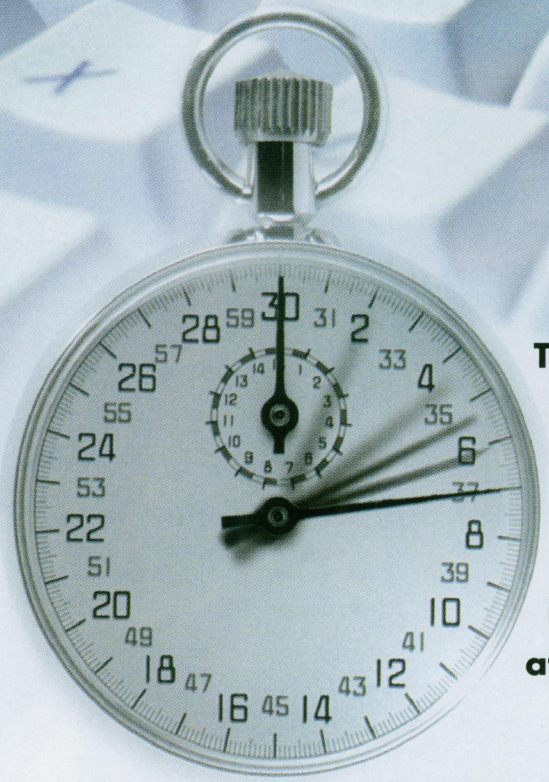
21. Bayley N. Manual for the Bayley Scales of Infant Development. New York:The Psychological Corporation, 1969.
22. McCarthy D. McCarthy Scales of Children's Abilities. New York:The Psychological Corporation, 1972.
23. Zimmerman I, Steiner V, Pond R. Preschool Language Scale. Rev ed. Columbus, OH:CE Merrill, 1979.
24. Woodcock R, Johnson M. Woodcock-Johnson Tests of Achievement. Allen, TX:DLM, 1989.
25. Koppitz E. The Bender Gestalt Test for Young Children. London, England:Grune & Stratton, 1963.
26. Achenbach T. Manual for the Child Behavior Checklist and 1991 Child Behavior Profile. Burlington, VT:University of Vermont Department of Psychiatry, 1991.
27. Raven J. Standard Progressive Matrices. Cambridge, England:HK Lewis, 1958.
28. Caldwell B, Bradley R. Home Observation of Measurement of the Environment. Little Rock, AR:University of Arkansas, 1984.
29. Cox DR, Oakes D. Analysis of Survival Data (Cox DR, Hinkley DV, eds). New York:Chapman and Hall, 1984.
30. Crump K, Van Landingham C. Bench_C: A Fortran Program to Calculate Benchmark Doses from Continuous Data. Ruston, LA:The K.S. Crump Group, 1995.
31. Gaylor DW, Slikker W. Risk assessment for neurotoxic effects. Neurotoxicology 11:211–218 (1990).
32. Crump K, Viren J, Silvers A, Clewell H, Gearhart J, Shipp A. Reanalysis of dose-response data from the Iraqi methylmercury poisoning episode. Risk Anal 15:523–532 (1995).
33. Crump K. On summarizing group exposures: is an arithmetic mean or a geometric mean more appropriate? Risk Anal 18:293–297 (1998).

ATTENTION RESEARCHERS!

Publish faster with *EHP*

Articles Online First

ehis
Environmental Health Information Service



The pace of scientific research is increasing all the time. *EHP* meets the demand for faster publication—all research articles accepted for publication in *EHP* are published online first through the Environmental Health Information Service at http://ehis.niehs.nih.gov/.

With *EHP* Articles Online First you can:

- publish your research months earlier than the print version and ahead of your competition
- precisely establish publication priority to a specific day
- stay informed on the very latest developments in environmental health science

Each online article is assigned a unique and permanent citation that includes volume, page numbers, online publication date, and specific URL that follows the paper into print publication.
For more information, refer to the Instructions to Authors located in the July issue of *EHP* and on the EHIS.

Visit us today! **http://ehis.niehs.nih.gov/**