Relation of Prenatal Phenylalanine Exposure to Infant and Childhood Cognitive Outcomes: Results From the International Maternal PKU Collaborative Study

Keith F. Widaman, PhD*, and Colleen Azen, MS‡

ABSTRACT. *Objective.* **The primary aims of this study were to model the form of the relation between prenatal exposure to phenylalanine (Phe) and measures of offspring intellectual development and to estimate the developmental relations of maternal demographic, pregnancy-related, and perinatal variables on offspring intelligence during infancy and childhood.**

Methods. **The participants were the 413 children and their mothers from the International Maternal PKU Collaborative Study.**

Results. **Results supported a nonlinear relation between prenatal Phe exposure and offspring cognitive outcomes, with damage to the developing fetus if average Phe levels are above approximately 360 μmol/L. Moreover, prenatal Phe exposure had a strong effect on offspring outcomes at 1 year of age and was the only one of the background, pregnancy-related, or perinatal variables to influence directly offspring outcomes at 2, 4, and 7 years of age.**

Conclusion. **The present study was able to document the importance of prenatal exposure to Phe for predicting offspring cognitive outcomes in the presence of other predictors of these outcomes.** ***Pediatrics*** **2003;112:1537–1543;** ***prenatal effects, phenylalanine, PKU, maternal PKU, intelligence, teratogenic effects, socioeconomic status, maternal intelligence, maternal age, pregnancy, birth outcomes, infancy, childhood, regression, spline models, structural equation models, longitudinal study.***

ABBREVIATIONS. Phe, phenylalanine; SES, socioeconomic status; PKU, phenylketonuria; SD, standard deviation; PAH, phenylalanine hydroxylase; MDI, Mental Development Index; PDI, Psychomotor Development Index; GCI, General Cognitive Index; WISC-R, Wechsler Intelligence Scale for Children–Revised.

Despite considerable research demonstrating detrimental effects of prenatal phenylalanine (Phe) exposure,[1] much remains to be determined with regard to the effects of prenatal exposure to Phe on offspring developmental outcomes. One key question is the form of the relation between prenatal Phe exposure and offspring outcomes. If this relation is linear, then any prenatal exposure to Phe might cause damage to the developing fetus and thus be teratogenic, although Phe exposure might have to reach some specifiable level (eg, 600 μmol/L) before clinically significant problems can be documented. Conversely, if the relation is nonlinear, then different functional forms could be considered, and critical thresholds might be identified beyond which teratogenic effects occur.

One candidate model for the relation between prenatal Phe exposure and offspring outcomes is a 2-piece linear spline, which is characterized by 2 linear functions that intersect at a "knot" point. The underlying assumption of this model is that the relation between the predictor and outcome variables is a linear function with a given slope below the knot point and a linear function with a different slope above the knot point.

A 2-piece linear spline model with 0 slope below the knot point and negative slope above the knot point is consistent with a threshold of prenatal Phe exposure; although hypothesized by many researchers, there is considerable variability in estimates of threshold values required for teratogenic effects to occur. For example, Perry et al[2] argued that maternal Phe levels >600 μmol/L were likely to be necessary before damage to the fetus occurs, but both Johnson[3] and Hsia[4] claimed that even higher levels of maternal Phe might be required to confirm a teratogenic effect of exposure. Consistent with this latter claim, Levy and Waisbren[5] found in a sample of 53 offspring that mental retardation occurred, with 1 exception, only when the mother's Phe level exceeded 1080 μmol/L. However, the survey by Lenke and Levy[1] documented mental retardation in 21% of offspring of mothers with Phe levels between 180 and 600 μmol/L.

A second important question involves the relative impact of prenatal exposure to Phe in the context of other variables. One of the most common correlates of offspring cognitive outcomes in infancy and childhood is parental intelligence,[6] because there are both genetic and environmental contributions by natural parents to their offspring that are thought to influence intelligence. Recent reviews of the behavioral genetic literature[7] conclude that genetic sources explain more than half of the variance in intelligence and that shared environmental factors explain small amounts of variance. In addition to parental intelligence, parental socioeconomic status (SES)[8] and parental education[9] are strong correlates of offspring intelligence throughout infancy, childhood, and ad-

From the *Department of Psychology, University of California at Davis, Davis, California; and ‡Children's Hospital Los Angeles, Los Angeles, California.

Reprint requests to (K.F.W.) Department of Psychology, University of California, One Shields Ave, Davis, CA 95616-8686. E-mail: kfwidaman@ucdavis.edu

PEDIATRICS (ISSN 0031 4005).

olescence. Advocates of genetic explanations of intelligence often argue that these presumed effects of SES and education of the parents on childhood outcomes are largely indirect effects of genetic sources of variance: parents with more favorable genetic endowments not only have higher intelligence but also have higher socioeconomic and education levels and provide more stimulating home environments.

Prenatal exposure to Phe is yet another factor that may influence offspring cognitive outcomes. Most research on maternal phenylketonuria (PKU) has focused on the relationship between maternal blood Phe levels during pregnancy and child outcomes, sometimes adjusting for other treatment and/or demographic factors, but has failed to estimate these effects within larger or more inclusive models for child developmental outcomes. Because mothers with higher prenatal Phe levels may have lower levels of intelligence, SES, and education, it is possible that maternal demographic factors may be more important than Phe levels during pregnancy when predicting offspring cognitive outcomes.

Thus, 2 general goals motivated the current study. The first goal was to determine the relative utility of linear and nonlinear regression models for representing the relation between prenatal Phe exposure and offspring cognitive outcomes. The second goal was to examine the relative effects of prenatal Phe exposure on offspring cognitive outcomes in the context of a range of maternal background variables, pregnancy-related variables, and indicators of perinatal status.

METHODS

Participants

The study sample consisted of 572 pregnancies, 412 of which resulted in the live birth of 416 offspring. Because 3 offspring were not entered into the follow-up study as a result of a diagnosis of PKU ($n = 2$) or parental refusal ($n = 1$), the final offspring sample numbered 413. Offspring were evaluated neonatally and then at 1, 2, 4, and 7 years of age.

Measures

Descriptive statistics for variables used in these analyses, presented in Table 1, include number of patients with valid scores, mean, standard deviation (SD), minimum, and maximum.

Demographic, or Background, Variables

Five demographic, or background, variables on the mothers were used in these analyses. SES of the household in which the mother resided was assessed using the Hollingshead 2-factor index, based on education and occupation of the head of household. Hollingshead scores were left in continuous form, ranging from 11 to 77, with larger scores indicating lower SES. Maternal education was scored on a categorical scale ranging from 1 (postgraduate degree) to 7 (<7 years). Maternal age at conception ranged from 16 to 36 years, with a mean of 24.1 years.

Maternal intelligence, assessed by the Verbal IQ and Performance IQ scores from the Wechsler Adult Intelligence Scale–Revised[10] fell, on average, approximately 1 SD below the population mean for intelligence.

The final background variable, reflecting the severity of the genetic mutation on the phenylalanine hydroxylase (*PAH*) gene, was assessed by 2 variables: 1) assigned PHE level, the maximum blood Phe level on an unrestricted diet, and 2) a score developed by Güttler and Guldberg[11] reflecting the severity of both mutations on the *PAH* gene, using DNA identification of the mutation. The Güttler score can range from 2 to 16, with higher numbers reflecting milder mutations.

Pregnancy-Related Variables

Five pregnancy-related variables were used in analyses reported in this article. Weight gain during pregnancy was assessed using 2 measures: 1) total weight gain in pounds during pregnancy and 2) percentage of recommended weight gain achieved at term. The second pregnancy-related variable was variability in Phe levels during pregnancy, assessed by the SD of all Phe levels during the pregnancy for an individual. The third was Phe level, assessed using 3 indicators: 1) the average of all Phe levels recorded during the pregnancy, 2) the week during pregnancy when

TABLE 1. Descriptive Statistics on Study Variables

Variable	*N*	Mean	SD	Minimum	Maximum
Hollingshead SES score	389	51.3	14.7	11	77
Mother's education code	397	3.9	1.2	1	7
Mother's age (y)	412	24.1	4.1	16	36
Mother's Verbal IQ	376	85.6	12.7	22	129
Mother's Performance IQ	378	88.5	14.4	47	132
Mother's Full-scale IQ	379	85.9	13.7	40	130
Assigned Phe level (μmol/L)	412	1322	551	198	3066
Güttler score	247	4.0	2.6	2	12
Weight gain (lb)	366	29.0	14.1	−13	70
% recommended weight gain	373	136.9	88.7	86	455
SD of Phe (μmol/L)	408	165	90	6	468
Average Phe exposure (μmol/L)	412	494	269	78	1698
Week Phe <600 (μmol/L)	412	15.0	14.3	0	41.7
Week Phe <360 (μmol/L)	412	25.2	14.4	0	42.3
Weeks' gestation at term	412	39.2	1.9	28.7	43.3
Birth length (cm)	406	49.0	3.1	39.0	56.5
Birth weight (g)	411	3068	562.8	1389	4886
Birth head circumference (cm)	403	32.8	2.0	26.0	38.0
Bayley MDI at 1 y	283	100.1	20.3	49	151
Bayley PDI at 1 y	265	98.1	18.3	49	138
Bayley MDI at 2 y	230	97.3	22.9	49	151
Bayley PDI at 2 y	208	99.0	20.5	49	146
McCarthy GCI at 4 y	276	85.2	21.2	45	132
TOLD CSLQ at 4 y	255	86.1	15.8	34	144
WISC-R Verbal IQ at 7 y	284	92.1	22.4	40	142
WISC-R Performance IQ at 7 y	285	92.0	21.9	40	133

TOLD CSLQ indicates Test of Language Development Composite Spoken Language Quotient.

all remaining Phe levels were <600 μmol/L, and 3) the week during pregnancy when all subsequent Phe levels were <360 μmol/L.

The fourth pregnancy-related variable was duration of the pregnancy, which ranged from approximately 28 to 43 weeks' gestation. The final pregnancy-related variable was average daily protein intake during pregnancy.

Birth Variables

Three birth variables were included in the models: offspring length, weight, and head circumference.

Measures of Cognitive Status

The measures of offspring cognitive status used in the present study were obtained at ages 1, 2, 4, and 7 years. At ages 1 and 2 years, the Bayley Scales of Infant Development[12] were administered, yielding Mental Development Index (MDI) and Psychomotor Development Index (PDI) scores at each age. When children were 4 years of age, the McCarthy Scales of Children's Abilities[13] were administered, which yielded General Cognitive Index (GCI) scores, as was the Test of Language Development,[14] from which the Composite Spoken Language Quotient was used. Finally, at 7 years of age, the Wechsler Intelligence Scale for Children–Revised (WISC-R)[15] was used to assess intelligence, from which the Verbal IQ and Performance IQ scores were used.

Analytic Procedures

To model the relation between prenatal Phe exposure and offspring intellectual outcomes, we used the REG and NLIN procedures within the SAS/STAT Software, Version 8 of the SAS system for PCs (SAS Institute, Inc, Cary, NC) to fit linear and 2-piece linear spline models to the data. Ordinary least-squares estimation was used under REG, and the multivariant secant method of estimation (option DUD) was used under NLIN assuming a slope of 0 for the segment representing lower maternal Phe levels. Because it is not possible to formulate statistical tests of the difference in fit of linear and 2-piece spline models, the models were evaluated with regard to their squared multiple correlation and the meaningfulness of their parameter estimates and associated standard errors.

When fitting structural equation models to the prenatal, perinatal, and offspring data, the LISREL 8.52 program[16] was used. Because of the presence of missing data for some participants, we used full information maximum likelihood estimation of model parameters. Full information maximum likelihood estimation has been found to be an efficient and unbiased method of estimation in the presence of missing data that are missing at random and is one of the best alternatives to use when data do not meet the "missing at random" assumption. Model fit was evaluated using both the likelihood ratio χ^2 goodness of fit statistical test and practical fit indices, as described by Jöreskog and Sörbom.[16]

RESULTS

Form of the Phe Exposure: Offspring Cognitive Outcome Relation

Regression analyses presented in this section were performed using maternal average Phe exposure during pregnancy recorded in mg/dL (1 mg/dL ≈ 60 μmol/L). Relevant values are reported in Table 2 in mg/dL, in both systems of units.

Outcomes at 1 Year of Age

The results of fitting linear and 2-piece linear spline models to offspring Bayley MDI scores at 1 year of age are shown in Table 2. Both models fit well, explaining >28% of the variance in MDI scores. Moreover, all parameter estimates were statistically significant, falling at least 7 standard errors from 0, but the 2-piece spline model explained approximately 1.5% more variance than did the standard linear model, so the spline model is the preferred model on the basis of this criterion.

TABLE 2. Predicting Offspring Outcomes at 1, 2, 4, and 7 Years of Age From Measures of Maternal Average Phe Level During Pregnancy

Equation	*R*	R^2	*F*	*df*	Root Mean Square Error
Offspring outcomes at 1 y of age					
Linear MDI = 119.53 (2.12) − 2.47 (0.24) (AVEPHE)	0.531	0.282	109.09	1, 277	17.10
Spline MDI = 108.11 (1.57) − 3.08 (0.36) [max (AVEPHE − KNOT, 0)], KNOT = 6.10 (0.85)	0.545	0.297	58.30	2, 276	16.96
Linear PDI = 113.45 (2.07) − 1.98 (0.24) (AVEPHE)	0.464	0.215	70.89	1, 259	16.10
Spline PDI = 103.01 (1.25) − 3.14 (0.50) [max (AVEPHE − KNOT, 0)], KNOT = 8.10 (0.92)	0.487	0.237	40.07	2, 258	15.91
Linear MDI = 121.33 (2.50) − 3.16 (0.28) (AVEPHE)	0.595	0.354	122.76	1, 224	18.58
Offspring outcomes at 2 y of age					
Spline MDI = 110.20 (2.47) − 3.30 (0.36) [max (AVEPHE − KNOT, 0)], KNOT = 3.90 (1.04)	0.592	0.350	60.04	2, 223	18.68
Linear PDI = 117.31 (2.50) − 2.48 (0.30) (AVEPHE)	0.508	0.258	70.38	1, 202	17.78
Spline PDI = 106.73 (1.92) − 3.01 (0.44) [max (AVEPHE − KNOT, 0)], KNOT = 5.60 (1.05)	0.515	0.265	36.23	2, 201	17.74
Offspring outcomes at 4 y of age					
Linear GCI = 109.23 (2.65) − 2.98 (0.29) (AVEPHE)	0.614	0.377	108.87	1, 180	16.16
Spline GCI = 98.52 (3.05) − 3.08 (0.33) [max (AVEPHE − KNOT, 0)], KNOT = 3.91 (1.22)	0.618	0.381	50.09	2, 179	16.14
Linear CSLQ = 103.87 (2.44) − 2.14 (0.27) (AVEPHE)	0.542	0.294	64.90	1, 156	13.39
Spline CSLQ = 95.38 (2.35) − 2.34 (0.33) [max (AVEPHE − KNOT, 0)], KNOT = 4.53 (1.32)	0.555	0.308	34.49	2, 155	13.30
Offspring outcomes at 7 y of age					
Linear VIQ = 112.65 (3.04) − 3.01 (0.33) (AVEPHE)	0.600	0.360	83.89	1, 149	17.35
Spline VIQ = 98.32 (2.16) − 4.09 (0.55) [max (AVEPHE − KNOT, 0)], KNOT = 6.80 (0.90)	0.623	0.388	46.97	2, 148	17.02
Linear PIQ = 114.16 (2.95) − 3.06 (0.32) (AVEPHE)	0.619	0.383	92.95	1, 150	16.86
Spline PIQ = 99.65 (2.07) − 4.25 (0.53) [max (AVEPHE − KNOT, 0)], KNOT = 6.79 (0.83)	0.652	0.425	54.97	2, 149	16.33

Tabled values in equations are parameter estimates, with standard errors in parentheses.
VIQ indicates Verbal IQ; PIQ, Performance IQ; AVEPHE, average PHE level in mg/dL (1 mg/dL ≅ μmol/L) during pregnancy; KNOT, knot point, or the point where the 2 linear functions join.

The parameters of the 2-piece linear spline model also seem to have a more reasonable interpretation than those from the standard linear model. The standard linear model had an intercept of 119.5 and a regression slope of −2.47. These estimates suggest that predicted offspring MDI scores at the lowest observed maternal Phe levels (1.3 mg/dL or 78 μmol/L) would approach 120 and that MDI scores decrease linearly approximately 2.5 points for every 1 mg/dL (60 μmol/L) increase in the mother's average Phe level. Thus, offspring of mothers with average Phe levels of 10 mg/dL (600 μmol/L) had a mean predicted MDI score of 95.3, or (120 − 2.47 × 10). The predicted MDI of 116 at the lowest maternal Phe is extremely unlikely in this sample, given the mean IQ of the mothers, which fell approximately 1 SD below the population mean.

In the 2-piece linear spline model, the intercept of 108.1 provides constant estimate of average offspring MDI scores for maternal Phe levels up to the estimated knot point of 6.10 mg/dL (366 μmol/L). Above the knot point, predicted MDI scores decrease more rapidly than in the linear model, 3.08 points for every additional 1 mg/dL (60 μmol/L) in maternal Phe. Thus, offspring of mothers with average Phe levels of 10 mg/dL (600 μmol/L) had a mean predicted MDI score of 96.1, or (108 − 3.08 × 3.9). The 2-piece linear spline model supports the conclusion that average Phe exposure below 6.1 mg/dL (366 μmol/L) is not teratogenic and has no measurable effect but that substantial fetal damage occurs when average maternal Phe exposure levels exceed 6.1 mg/dL (366 μmol/L), with greater damage as Phe levels rise higher above the knot point.

The results for the PDI at 1 year of age, shown in Table 2, are similar to those for the MDI. Although the linear model explained considerable variance in PDI scores, the spline model explained approximately 2% more variance, all parameter estimates in both equations were statistically significant, and the parameter estimates from the spline model were more interpretable. The estimated knot point for PDI scores was 8.1 mg/dL (486 μmol/L), suggesting that higher levels of Phe exposure are required to affect PDI scores negatively than was the case for MDI scores.

Outcomes at 2 Years of Age

Table 2 report results for Bayley scores at 2 years of age. The linear model for the MDI explained slightly more variance than did the 2-piece linear spline model, a surprising outcome because the latter model has 1 more parameter than does the former model. Despite having slightly superior fit, the standard linear model had less interpretable parameter estimates, with predicted MDI scores approaching 120 at the lowest maternal Phe levels (ie, the intercept of 121.3). The 2-piece spline model had predicted MDI scores of approximately 110 for offspring whose mothers maintained average Phe levels below the knot point of 3.9 mg/dL (234 μmol/L) and a decrease of 3.3 MDI score units for every 1 mg/dL (60 μmol/L) increase in average Phe exposure beyond the knot point.

For Bayley PDI scores at 2 years of age, the 2-piece linear spline model explained more variance than did the standard linear model and also had more interpretable parameter estimates. The knot point of 5.6 mg/dL (336 μmol/L) for PDI scores was larger than that for MDI scores, suggesting that higher levels of Phe exposure are required to affect PDI scores negatively than was the case for MDI scores, but note that the knot points at 2 years of age (3.9 for MDI and 5.6 for PDI) were >2 mg/dL (120 μmol/L) lower than the estimates at 1 year of age (6.1 for MDI and 8.1 for PDI).

Outcomes at 4 Years of Age

Table 2 provides results from modeling McCarthy GCI scores. The 2-piece linear spline fit the data only slightly better than the standard linear model, but the parameter estimates for the spline model were more interpretable. In particular, predicted GCI scores were close to the population mean for offspring of mothers with Phe levels below the knot point of 3.9 mg/dL (234 μmol/L), after which GCI scores decreased approximately 3.1 points for every 1 mg/dL (60 μmol/L) increase in average Phe exposure. The knot point for GCI at age 4 was virtually identical to the knot point for MDI scores at 2 years of age.

Table 2 also shows results for Composite Spoken Language Quotient scores. The results closely approximated those for the GCI, with a better fit for the spline model and a knot point estimate of 4.5 mg/dL (270 μmol/L), after which teratogenic effects of Phe exposure occur.

Outcomes at 7 Years of Age

Table 2 show results for modeling WISC-R IQ scores at 7 years of age. Mirroring previous results, the 2-piece linear spline fit the data better than did the standard linear regression, explaining almost 3% more variance. Although all parameter estimates in both equations were statistically significant, the parameter estimates from the 2-piece linear spline were more easily interpreted. Specifically, predicted Verbal IQ scores were close to the population mean for offspring whose mothers had average Phe levels below the knot point of 6.8 mg/dL (408 μmol/L), above which Verbal IQ scores fell >4 points for every 1 mg/dL (60 μmol/L) increase in Phe exposure.

The results for Performance IQ scores closely parallel those for the Verbal IQ, with the 2-piece linear spline explaining >4% more variance than the standard linear regression model. The estimates from the 2-piece linear spline model predicted offspring Performance IQ near the population mean for maternal Phe below the knot point at approximately 6.8 mg/dL (408 μmol/L) and an estimated decrease in IQ scores of 4.25 points for every 1 mg/dL (60 μmol/L) increase after the knot point.

Modeling Influences of Prenatal and Perinatal Variables on Offspring Intellectual Outcomes

To model the relationship of maternal background, pregnancy-related, and birth variables to offspring cognitive outcomes, we formulated a large

structural equation model that embodied 4 sets of variables. The first set of variables consisted of 5 background latent variables: 1) mother's SES, with Hollingshead scores as measured variable; 2) mother's education, with education ratings as its indicator; 3) mother's IQ, with the mother's Verbal IQ and Performance IQ scores from the Wechsler Adult Intelligence Scale–Revised as indicators; 4) mother's age, with chronological age at conception of child as indicator; and 5) severity of mutation, with assigned PHE level and Güttler score as the 2 indicators. In this analysis, the "direction" of variables for which higher values were less desirable (SES score and education code) was mathematically transformed so that higher values indicated better maternal status.

Five pregnancy-related latent variables were included: 1) weight gain, with total weight gain and percentage of recommended weight gain as indicators; 2) PHE variability, with SD of Phe levels during pregnancy as indicator; 3) PHE level, with average Phe level during pregnancy, weeks' gestation after which all subsequent Phe levels were <600 μmol/L, and weeks' gestation after which all subsequent Phe levels were <360 μmol/L as the 3 indicators; 4) weeks' gestation, with number of weeks' gestation at birth as indicator; and 5) protein intake, with average daily protein intake (in grams) as indicator.

Two birth-related latent variables were specified: 1) birth size, with birth weight and birth length as its 2 indicators, and 2) birth head circumference, with measured head circumference at birth as indicator.

Finally, 4 offspring cognitive outcome latent variables were specified: 1) Bayley MDI at 1 year; 2) Bayley MDI at 2 years; 3) McCarthy GCI at 4 years; and 4) WISC IQ, with Verbal IQ and Performance IQ as its 2 indicators. The first 3 of the offspring cognitive outcome latent variables were single-indicator latent variables.

Given limitations of space, the modeling of the relations among the preceding variables cannot be described in detail. Three latent variables—weight gain, protein intake, and birth size—were excluded from the final model because the background latent variables had negligible effects on these variables and because the variables had negligible or 0 effects on offspring cognitive outcomes. In the final model, all single-indicator latent variables were assumed to be error-free, and indicators on multiple-indicator latent variables had moderate to high loadings (0.60–0.95, in standardized metric) on their respective latent variables.

The final structural equation model exhibited adequate levels of fit, with a root mean square error of approximation[16] of 0.055. The final structural model, showing all directed paths in the model, is shown in Fig 1. All path coefficients shown in Fig 1 are statistically significant ($P < .05$), and all path coefficients

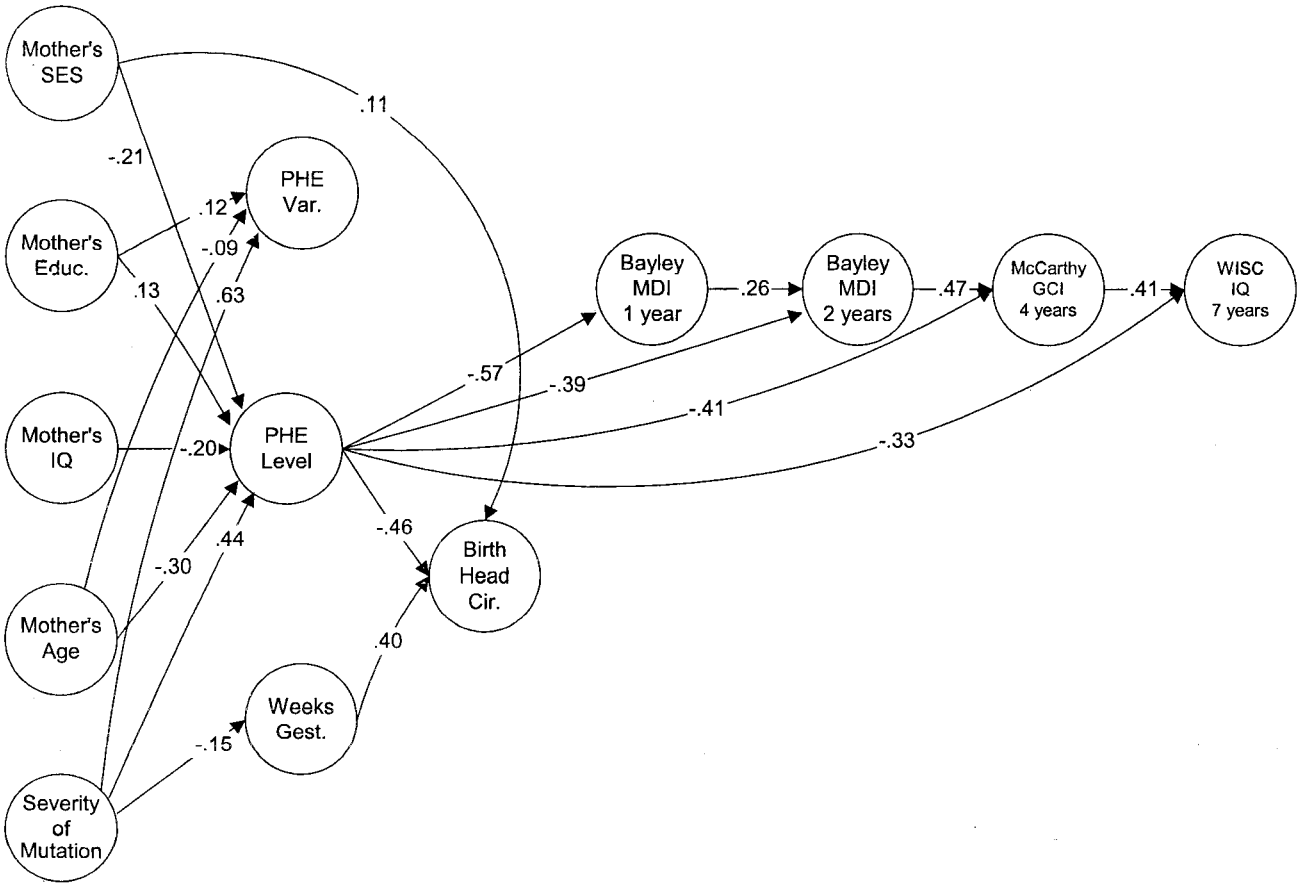

Fig 1. Final structural equation model representing relations among the maternal background, pregnancy-related, birth-related, and offspring cognitive outcome variables.

with absolute magnitude >0.20 were significant past the $P < .01$ level.

First, we discuss effects of maternal background variables on the pregnancy-related latent variables. As shown in Fig 1, PHE variability was influenced by 3 background latent variables (mother's education, mother's age, and severity of mutation). The only sizable effect was that of severity of mutation, with mothers with more severe mutations exhibiting greater variability in their Phe levels during pregnancy ($\beta = 0.63$). PHE level was influenced by all 5 of the background latent variables, and 4 of the 5 path coefficients were of moderate size (>0.20 in absolute magnitude). Mothers with higher levels of SES and intelligence had lower average Phe levels ($\beta = -0.21$ and $\beta = -0.20$, respectively). The 2 strongest effects were those of mother's age and the severity of her mutations, with older mothers having lower average Phe levels ($\beta = -0.30$) and mothers with more severe mutations having higher average Phe levels ($\beta = 0.44$). As with PHE variability, the strongest effect on PHE level was made by the severity of mutation latent variable. The final pregnancy-related latent variable still in the model, weeks' gestation, was influenced only by severity of mutation, with mothers with more severe mutations having somewhat shorter pregnancies, although this effect was modest in magnitude ($\beta = -0.15$).

The single birth latent variable remaining in the model was birth head circumference. Three variables had significant effects on head circumference, although the effect of mother's SES was of minor importance ($\beta = 0.11$). The 2 strong effects were those of PHE level during pregnancy ($\beta = -0.46$), with higher PHE levels leading to smaller head circumference, and weeks' gestation ($\beta = 0.40$), with longer pregnancies leading to newborns with larger head circumferences.

Turning to the 4 offspring cognitive outcome latent variables, we first specified autoregressive paths between these variables. That is, Bayley MDI at 1 year was allowed to influence Bayley MDI at 2 years, Bayley MDI at 2 years influenced McCarthy GCI at 4 years, and McCarthy GCI at 4 years influenced WISC-R IQ at 7 years. We then tested the effects of the maternal background, pregnancy-related, and birth-related latent variables on the cognitive outcome latent variables. Only PHE level during pregnancy had significant direct effects on the offspring cognitive outcome latent variables. That is, once PHE level during pregnancy was allowed to predict the cognitive outcome latent variables, none of the remaining potential predictors—not even mother's IQ or birth head circumference—had significant effects on the cognitive outcome latent variables. Moreover, the effects of PHE level on the cognitive outcomes were pervasive. As shown in Fig 1, PHE level had a strong effect on Bayley MDI at 1 year ($\beta = -0.57$) and also had strong direct effects on Bayley MDI at 2 years ($\beta = -0.39$), McCarthy GCI at 4 years ($\beta = -0.41$), and WISC-R IQ at 7 years ($\beta = -0.39$), despite the presence of autoregressive paths (ranging between 0.26 and 0.47) between the successive cognitive outcome measures.

DISCUSSION

There were 2 aims of the current study. The first was to determine the form of the relationship between prenatal Phe exposure and infant and childhood cognitive outcomes. Even at 1 year of age, the 2-piece linear spline model clearly fit the Bayley MDI and PDI data better than did a standard linear regression model. At 2 years of age, the linear and 2-piece linear spline models had approximately equal levels of fit to MDI and PDI scores, but the 2-piece linear spline model consistently had more easily interpretable parameter estimates. Then, for offspring outcomes at 4 and 7 years of age, the 2-piece linear spline models had both higher levels of explained variance and more easily interpreted parameter estimates than did the standard linear models. Thus, the form of the relation between prenatal Phe exposure and offspring cognitive outcomes seems to be nonlinear, with no damage to the developing fetus until exposure passes a critical threshold level, estimated as the knot point in the 2-piece spline model.

The knot point in the 2-piece linear spline model is a key parameter, as exposure to levels falling below the knot point presumably have no teratogenic effects whereas exposure to levels above the knot point result in damage to the offspring. The estimates of the knot point varied somewhat across offspring age levels and cognitive outcomes, ranging between 234 and 486 μmol/L. Averaging across the several estimates of the knot point leads to an estimate of 330 to 360 μmol/L as the best estimate of the critical threshold of Phe exposure beyond which damage to the offspring occurs. A critical threshold in the range from 330 to 360 μmol/L is consistent with the data presented by Lenke and Levy,[1] which showed elevated rates of mental retardation even in offspring with maternal Phe levels between 180 and 600 μmol/L, and is inconsistent with the much higher threshold values discussed by many researchers.[2,5]

The second goal was to estimate the effects of prenatal exposure to Phe within the context of a model that included an array of maternal background, pregnancy-related, and birth-related variables. The results of this modeling were concise. The severity of the mutation on the mother's *PAH* gene had stronger influences than the other background variables on pregnancy-related variables, including on both level and variability of maternal Phe levels during pregnancy. Thus, mothers with more severe mutations had higher Phe levels during pregnancy and exhibited greater variability in their Phe levels.

Of the pregnancy-related latent variables, only PHE level had notable effects on birth-related and offspring cognitive outcome variables, and all of these effects were large. PHE level had a large effect on birth head circumference and then had moderate to large direct effects on all 4 of the cognitive outcome latent variables. Importantly, PHE level mediated the effects of maternal background variables on offspring cognitive outcomes. That is, all 5 of the maternal background variables had significant influences on PHE level, and PHE level then significantly

affected all remaining birth-related and cognitive outcomes. Moreover, with 1 minor exception (the effect of mother's SES on birth head circumference), the maternal background variables had no significant direct effects on the birth-related and cognitive outcomes once the effects of PHE level on these latent variables were estimated.

Taken together, these results point to the crucial nature of Phe level during pregnancy with regard to offspring cognitive outcomes. The model suggests that the influence of the *PAH* mutations on offspring cognitive outcomes could be prevented if Phe levels during pregnancy were controlled at low levels. If all mothers with PKU could keep their average Phe levels during pregnancy below the crucial threshold value of 330 to 360 μmol/L, then teratogenic effects of Phe exposure during pregnancy would likely be negligible or nonexistent, and the offspring of these mothers would have every expectation of having cognitive skills within the normal range.

ACKNOWLEDGMENTS

The present work was supported in part by contract N01-HD-2-3148 from the National Institute of Child Health and Human Development and by intramural grants from the University of California at Davis.

REFERENCES

1. Lenke RR, Levy HL. Maternal phenylketonuria and hyperphenylalaninemia: an international survey of the outcome of untreated and treated pregnancies. *N Engl J Med*. 1980;303:1202–1208
2. Perry TL, Hansen S, Tischler B, Richards FM, Sokol M. Unrecognized adult phenylketonuria: implications for obstetrics and gynecology. *N Engl J Med*. 1973;289:395–398
3. Johnson CF. Phenylketonuria and the obstetrician. *Obstet Gynecol*. 1972; 39:942–947
4. Hsia DY-Y. Phenylketonuria and its variants. *Progr Med Genet*. 1970;7: 29–68
5. Levy HL, Waisbren SE. Effects of untreated maternal phenylketonuria and hyperphenylalaninemia on the fetus. *N Engl J Med*. 1983;309: 1269–1274
6. Sattler JM. *Assessment of Children: Cognitive Applications*. 4th ed. San Diego, CA: Jerome M. Sattler, Publisher; 2001
7. McGue M, Bouchard TJ Jr. Genetic and environmental influences on human behavioral differences. *Annu Rev Neurosc*. 1998;21:1–24
8. Bradley RH, Corwyn RF. Socioeconomic status and child development. *Annu Rev Psychol*. 2002;53:371–399
9. Broman S, Nichols PL, Shaughnessy P, Kennedy W. *Retardation in Young Children: A Developmental Study of Cognitive Deficit*. Hillsdale, NJ: Erlbaum; 1987
10. Wechsler D. *Wechsler Adult Intelligence Scale–Revised*. San Antonio, TX: The Psychological Corporation; 1981
11. Guldberg P, Rey F, Zschocke J, et al. A European multicenter study of phenylalanine hydroxylase 6 deficiency: classification of 105 mutations and general system for genotype-based prediction of metabolic phenotype. *Am J Hum Genet*. 1998;63:7179
12. Bayley N. *Bayley Scales of Infant Development*. 2nd ed. San Antonio, TX: The Psychological Corporation; 1993
13. McCarthy D. *Manual for the McCarthy Scales of Children's Abilities*. New York, NY: The Psychological Corporation; 1972
14. Newcomer PL, Hammill DD. *Test of Language Development–Primary*. 2nd ed. Austin, TX: Pro-Ed; 1991
15. Wechsler D. *Wechsler Intelligence Scale for Children–Revised*. San Antonio, TX: The Psychological Corporation; 1974
16. Jöreskog KG, Sörbom D. *LISREL 8: User's Reference Guide*. 2nd ed. Chicago, IL: Scientific Software International; 1999

Relation of Prenatal Phenylalanine Exposure to Infant and Childhood Cognitive Outcomes: Results From the International Maternal PKU Collaborative Study

Keith F. Widaman and Colleen Azen

Pediatrics 2003;112;1537

PEDIATRICS®
OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Relation of Prenatal Phenylalanine Exposure to Infant and Childhood Cognitive Outcomes: Results From the International Maternal PKU Collaborative Study

Keith F. Widaman and Colleen Azen

Pediatrics 2003;112;1537