Prenatal methylmercury exposure from ocean fish consumption in the Seychelles child development study

Gary J Myers, Philip W Davidson, Christopher Cox, Conrad F Shamlaye, Donna Palumbo, Elsa Cernichiari, Jean Sloane-Reeves, Gregory E Wilding, James Kost, Li-Shan Huang, Thomas W Clarkson

Summary

Introduction Exposure to methylmercury (MeHg) before birth can adversely affect children’s neurodevelopment. The most common form of prenatal exposure is maternal fish consumption, but whether such exposure harms the fetus is unknown. We aimed to identify adverse neurodevelopmental effects in a fish-consuming population.

Methods We investigated 779 mother-infant pairs residing in the Republic of Seychelles. Mothers reported consuming fish on average 12 meals per week. Fish in Seychelles contain much the same concentrations of MeHg as commercial ocean fish elsewhere. Prenatal MeHg exposure was determined from maternal hair growing during pregnancy. We assessed neurocognitive, language, memory, motor, perceptual-motor, and behavioural functions in children at age 9 years. The association between prenatal MeHg exposure and the primary endpoints was investigated with multiple linear regression with adjustment for covariates that affect child development.

Findings Mean prenatal MeHg exposure was 6.9 parts per million (SD 4.5ppm). Only two endpoints were associated with prenatal MeHg exposure. Increased exposure was associated with decreased performance in the grooved pegboard using the non-dominant hand in males and improved scores in the hyperactivity index of the Conner’s teacher rating scale. Covariates affecting child development were appropriately associated with endpoints.

Interpretation These data do not support the hypothesis that there is a neurodevelopmental risk from prenatal MeHg exposure resulting solely from ocean fish consumption.

Lancet 2003; 361: 1686–92
See Commentary page 1667

Methods

Participants

In 1989–90 we enrolled 779 mother-child pairs (about 50% of live births during that period), when the children were 6 months old. We excluded mothers and children with disorders highly associated with adverse neurodevelopment such as traumatic brain injury, meningitis, epilepsy, and severe neonatal illnesses. Although these disorders have been associated with overt and subtle neurodevelopmental problems, no data exist to suggest they are associated with MeHg exposure. The 44 exclusions through 66 months of age have been reported. We subsequently excluded 18 children for...
Concentrations of Hg are expressed in µg/g, where 1 µg/g = 1 part per million (ppm) in hair.

We assessed neurocognitive, language, memory, motor, perceptual-motor, and behavioural functions. Our tests included overall and domain-specific items, covering neurodevelopmental domains associated with prenatal MeHg exposure and included most of the specific tests used in previous studies. Individual tests measured intelligence (the Wechsler intelligence scale for children III [WISC III] full-scale IQ); learning and achievement (the Woodcock-Johnson test of achievement, letter-word recognition, and applied problems subtests and the California verbal learning test); memory (the visual memory subtest of the wide-range assessment of memory and learning); motor functions (finger tapping, trailmaking, grooved pegboard, and most of the Bruininks-Oseretsky test of motor proficiency); language (Boston naming test); visual-motor integration (the Beery-Buktenica developmental test of visual motor integration and a test of haptic matching); and sustained attention (Connor’s continuous performance test). We assessed behaviour with the Connor’s teacher rating scale and the parent-child behaviour checklist.

A team of three Seychellois child health and development professionals (a senior nurse, a child psychologist, and a special educator) assessed the children.

They received extensive training in child development and psychometric assessment procedures at the University of Rochester before the assessments. All personnel working in Seychelles were unaware of the MeHg exposure from the start of the study and no individual MeHg concentrations have been shared with families, clinical investigators, or anyone in Seychelles.

We investigated test reliability among testers and between each tester and a psychologist (PWD or DP). Pair-wise intertester reliability was assessed once a week by having a child’s performance during testing scored by two team members simultaneously. We investigated gold-standard reliability by on-site simultaneous scoring of about 5% of test sessions by one of the psychologists.

The cohort was initially tested between February, 1997, and November, 1998. Every child was seen twice about 1 month apart. The sessions lasted about 3 h. During the first session the caregiver completed a demographic questionnaire and the parental child behaviour checklist while the child was given the WISC III and had audiometry and tympanometry. All remaining tests were administered individually during the second session. Testing took place in a specially established child development centre, mostly in the morning, and the tests were given in the same sequence in each session.

Table 2: Results of neurodevelopment tests for cognition and achievement by prenatal exposure

<table>
<thead>
<tr>
<th>WJ test</th>
<th>Data by MeHg in maternal hair (mean [SE])</th>
<th>Regression coefficient (SE)*</th>
<th>p</th>
<th>95% CI for a 10 µg/g change</th>
</tr>
</thead>
<tbody>
<tr>
<td>LW recognition</td>
<td>90–110</td>
<td>131-7 (40-3)</td>
<td>-0.012 (0.046)</td>
<td>0.79</td>
</tr>
<tr>
<td>Applied problems</td>
<td>90–110</td>
<td>95-4 (15-5)</td>
<td>0.059 (0.015)</td>
<td>0.71</td>
</tr>
<tr>
<td>Higher scores in all tests indicate better performance. WISC III-Wechsler intelligence scale for children version III. CVLT-California verbal learning test. BNT-Boston naming test. WJ-Weech-Personson-Johnson test of achievement. LW Recognition-letter word recognition. *Data from prenatal mercury term in the multiple regression models. †Henderson early learning process scale. ‡HOME-home observation for measurement of the environment. §K-Bit=Kaufman brief intelligence test to determine caregiver intelligence.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the Kaufman brief intelligence test to determine caregiver environment. They also completed the matrices subtest of measure the quality of stimulation in the home were recalled in 1999–2000 and given the family resource were the child’s biological mother). Primary caregivers least 5 days per week hosted the visit (93% of caregivers the environment. The child’s primary caregiver, defined was made to administer the Caldwell-Bradley preschool was normal range). Normal Overall data Data by MeHg in maternal hair (mean [SE]) Regression coefficient (SE)* p CI for a 10 µg/g change‡

| VMI | 100 (15) | 96-0 (11-7) | 95-1 (12-2) | 98-6 (0-8) | 97-9 (0-9) | 95-7 (1-2) | 96-6 (1-4) | -0-010 (0-012) | 0-03 | -2-4 to 2-2 |
| Brunnings-Oseretsky | -- | 44-6 (6-1) | 44-3 (0-6) | 45-0 (0-5) | 45-0 (0-6) | 45-3 (0-6) | -0-093 (0-056) | 0-10 | -0-2 to 2-0 |
| Haptic discrimination test (total correct out of 10) | -- | 4-1 (1-8) | 4-2 (0-2) | 4-1 (0-1) | 4-2 (0-2) | 4-2 (0-2) | -0-010 (0-018) | 0-60 | -0-5 to 0-3 |

| Grooved pegboard time (s) | Dominant hand‡ | 74 (15) | 91-8 (20-5) | 89-5 (79-100) | 89-0 (80-100) | 87-7 (78-101) | 89-0 (78-99) | 3-3×10–4 (1-9×10–4) | 0-08§ | 91-4 to 98-1 |
| | Non-dominant hand¶ | 81 (16) | 100-1 (18-9) | 95-5 (87-104) | 93-4 (84-105) | 108-3 (93-113) | 98-3 (89-110) | 6-5×10–4 (2-5×10–4) | 0-01 | 101-7 to 112-9 |
| | Female | 108-2 (29-8) | 106-0 (95-124) | 101-0 (90-120) | 101-7 (89-120) | 101-7 (85-120) | 98-8 (90-113) | -2-5×10–4 (2-6×10–4) | 0-34 | 100-0 to 111-3 |
| | Male | 100-5 (20-5) | 86-9 (78-100) | 89-0 (79-100) | 87-7 (78-101) | 89-0 (78-99) | 3-3×10–4 (1-9×10–4) | 0-08§ | 91-4 to 98-1 |

| Trail making time (s) | A‡ | 25 (9) | 33-7 (17-1) | 30-0 (24-40) | 29-2 (22-40) | 29-2 (23-39) | 31-2 (25-39) | 31-2 (23-38) | 0-0037 (0-0038) | 0-33 | 32-5 to 37-6 |
| | B§ | 55 (19) | 81-5 (49-6) | 67-6 (49-101) | 63-8 (48-90) | 65-8 (52-101) | 65-8 (49-91) | 62-6 (47-89) | 0-0067 (0-0050) | 0-17 | 79-1 to 96-0 |

| Finger tapping | Dominant hand | M 40 (5) | 34-0 (5-7) | 34-3 (0-5) | 33-8 (0-5) | 34-7 (0-6) | 33-7 (0-5) | -0-050 (0-053) | 0-34 | -1-5 to 0-5 |
| | Non-dominant hand | M 35 (5) | 30-0 (4-6) | 29-5 (0-4) | 30-0 (0-4) | 30-1 (0-4) | 31-0 (5-5) | 29-8 (0-4) | 0-016 (0-041) | 0-69 | |

| WRAML design memory | 10 (3) | 7-7 (2-9) | 7-7 (0-2) | 7-6 (0-2) | 7-9 (0-3) | 7-7 (0-3) | 7-8 (0-3) | -0-021 (0-029) | 0-48 | -0-8 to 0-4 |

Higher scores indicate better performance except for Grooved Pegboard and Trailmaking where higher scores indicate poorer performance. VMI=visual motor integration. WRAML=wide range assessment of memory and learning. *Data from prenatal mercury term in the multiple regression models.

Table 3: Neurodevelopmental tests for motor, perceptual motor and memory by prenatal exposure

When the cohort was 42–56 months of age, a home visit was made to administer the Caldwell-Bradley preschool version of the home observation for measurement of the environment. The child’s primary caregiver, defined as the family member with whom the child lived for at least 5 days per week hosted the visit (93% of caregivers were the child’s biological mother). Primary caregivers were recalled in 1999–2000 and given the family resource scale and the Henderson early learning process scale to measure the quality of stimulation in the home environment. They also completed the matrices subtest of the Kaufman brief intelligence test to determine caregiver intelligence.

The covariates used in the analysis are shown in table 1. They were selected a priori for their known effect on child development and were expected to provide an index of the effectiveness of the assessments. The Hollingshead four-factor socioeconomic score was calculated with a list of Seychellois employment codes. Recent postnatal MeHg exposure was included since it was associated with outcomes in the 66-month assessments. It was measured in a 1-cm segment of hair closest to the scalp on a sample taken at the initial 9-year assessment, and represented about 1 month of recent exposure. The mean postnatal hair concentration was 0·1 µg/g (SD 3·5). Lead, polychlorinated biphenyls, and pesticides were not polychlorinated biphenyls, and pesticides were not.
we reran the model, dropping the non-significant
only the prenatal or postnatal interaction were significant,
significant, we report the model without interactions. If
reported for this model. If neither interaction was
first. If both interactions were significant then results are
significance level of 0.05, the results of that analysis were
both models was not significant at a two-tailed
included since measured concentrations in Seychelles
analysis so the y-axis is nonlinear.
with mercury concentrations in the multiple regression model. Lines are
shown for the model with outliers (dashed) and without outliers (solid). Outliers are indicated by larger symbols. Test scores were transformed for analysis so the y-axis is nonlinear.

The primary analysis included 21 endpoints. We used the
main score for most tests, but several measures yielded
more than one endpoint. For every endpoint, we did a
maximum of three linear-regression analyses for prenatal
MeHg exposure using all the covariates defined in table 1. All
analyses were done with the SAS system, version 8. Because differential effects on males and females have
been reported, every model was run first with and then
without a MeHg by sex interaction term for both prenatal
and recent postnatal exposure.10,18–20 If the overall test for
this report.

Every analysis included an assessment of residuals as a
check on the assumptions of normally distributed errors
with constant variance. If the assumptions seemed to be
violated, we used power transformations to stabilise the
variance and produce more normally distributed errors.
For every analysis, we assessed the model for statistical
outliers (scores with standardised residual values >3 or
<−3). All models with outliers were rerun without the
outliers and the results with and without outliers were
compared and are reported.

We also assessed every regression model for the effects
of influential points, identified by deleting each point in the data set individually from the analysis and
calculating the resulting standardised change in the
regression coefficient for prenatal MeHg exposure. The
regression analysis for all primary endpoints was repeated
without influential points to determine whether the
original results were dependent upon such points. The
final analysis included influential points that were not also
outliers.

Role of the funding source
The sponsors of the study approved the study design but
had no other involvement in the study design, data
collection, data analysis, data interpretation, or writing of
this report.

Results
The mean prenatal total MeHg exposure was 6.9 μg/g
(SD 4.5). The correlation coefficient between prenatal
and postnatal exposure was −0.08 (p=0.04).

Intraclass correlation coefficients were computed for
each of the 13 subtests of the WISC III. This test was
chosen for reliability computations since it was the most
difficult component to administer. Comparisons between
WISC-III subtest scores obtained by pairs of testers
(n=49) ranged from 0.90 to 1.00. Agreement between
each tester and one of the team psychologists (n=37)
ranged from 0.81 to 1.00.

The mean age at testing was 107 months (SD 4). For
data presentation, the mean (SD) for each endpoint was
computed by prenatal mercury exposure groupings. These
results are shown in table 2 (cognition and achievement
domains), table 3 (motor, perceptual-motor, and memory
domains), and table 4 (attention and behaviour domains).
In cases where the endpoint was transformed for analysis,
the median and quartiles have been included in the tables,
rather than the mean. Seychellois children’s scores on most endpoints compared favourably with US norms. The variability of the tests was as expected and they seemed to discriminate well among cohort children. The WISC III and the Boston naming test (table 2) were both affected by cultural variation, with lower means for Seychellois children than for US controls. However, the variability associated with these and other endpoints was consistent with test expectations, suggesting that all tests discriminated well among cohort children. Seychellois children did substantially better than the US norms on the Woodcock-Johnson letter-word recognition test, measuring scholastic achievement in reading.

Significant two-tailed overall model F statistics (p<0.05) resulted from the regression analysis for 20 of the 21 models with outliers removed. The squared multiple correlation coefficients for these models ranged between 5% and 24% with most values between 10 and 20%. Significant associations between prenatal MeHg exposure and performance were found for two endpoints, and both needed transformation for analysis. There was a significant decrease in performance on the grooved pegboard time for the non-dominant hand in males (table 3; figure 1). This task required an average of 100 s for cohort participants to complete and a 95% CI for a 10 µg/g change in exposure ranged from 1 to 13 s. The variability for this task within the reference population was 19 s, suggesting the effect was small. The association for the same task using the dominant hand was not significant. There was a significant improvement of the hyperactivity index of the Connor’s teacher-rating scale (table 4; figure 2) as prenatal MeHg increased. A 10 µg/g increase in exposure would result in a drop of between 1 and 6 points (95% CI) in the hyperactivity index.

| Caregiver Maternal Child’s Home Socioeconomic score | IQ | Maternal Chid’s Home Age | Overall Low Med High | Overall Unskl Semi-skill Skill Professional |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WISC III full scale IQ | 0.14 | 0.20 | -2.08 | -1.61 | -0.40 | 2.02 | -2.61 | 0.28 | -0.20 | -0.20 | 2.53 |
| CVLT Short delay | 0.001 | -0.01 | -0.61 | -0.11 | 0.05 | 0.06 | -0.17 | -0.004 | -0.05 | 0.22 |
| CVLT Long delay | 0.0001 | -0.002 | -0.57 | -0.05 | -0.04 | 0.09 | -0.26 | 0.03 | 0.10 | 0.13 |
| Boston naming test total score | 0.03 | 0.06 | 1.45 | -1.10 | 0.13 | 0.97 | -0.94 | -0.34 | 0.22 | 1.06 |
| WJ test Letter-word recognition | 0.44 | 0.64 | 13.7 | 1.37 | -7.60 | 6.23 | -3.69 | -0.34 | 1.08 | 2.95 |
| WJ test Applied problems | 0.24 | 0.20 | 0.96 | -1.21 | -0.79 | 7.00 | -3.62 | 0.66 | -1.03 | 3.98 |
| WJ test VMI | 0.11 | 0.03 | -3.89 | -1.03 | -0.68 | 1.71 | -1.33 | 0.98 | 0.86 | -0.52 |
| B-O test of motor development Haptic discrimination test | 0.03 | -0.01 | 2.42 | -0.004 | -0.36 | 0.36 | -0.14 | 0.43 | 0.32 | 0.26 |
| Trailmaking A | -0.003 | -0.004 | -0.30 | 0.03 | -0.02 | -0.01 | 0.0001 | 0.01 | 0.01 | -0.02 |
| Finger tapping Dominant hand | 0.02 | 0.06 | 3.33 | -0.76 | 0.09 | 0.67 | -0.31 | -0.30 | 0.45 | 0.16 |
| WRAML design memory | 0.03 | 0.02 | -0.22 | -0.43 | 0.05 | 0.38 | -0.25 | -0.03 | 0.08 | 0.19 |
| CPT Hit reaction | 0.13 | 0.20 | 0.99 | -2.61 | 0.92 | 1.69 | 0.33 | 0.10 | -0.56 | 0.13 |
| CTRs hyperactivity index | -0.001 | 0.0003 | 0.05 | -0.01 | 0.01 | 0.04 | 0.02 | 0.02 | -0.01 | -0.02 |

Table 5: Social and environmental covariate effects
Seychellois children were functioning at the upper limit of the normal range for this test.

The large number of endpoints raised concern about multiplicity and so we investigated the actual distributions of the p values for the prenatal exposure for consistency with the overall null hypothesis of no association. If the overall null hypothesis were true, the p values should be uniformly distributed from zero to one. We assessed this graphically by plotting the p values against idealised values from the uniform distribution. The distribution of p values was consistent with the expected values, and provides support for the overall null hypothesis of no association (figure 3). For comparison we plotted the p values for the home observation for measurement of the environment, a covariate from the same series of analyses that was associated with the endpoints.

Table 5 shows the associations between endpoints and selected social and environmental covariates. These factors have well established associations with child development and were expected to provide an index of the effectiveness of the assessments in ascertaining developmental status of the children. The data suggest that the effects of these factors on the endpoints were consistent with established associations. For example, socioeconomic score, early home environment scores, and maternal IQ were consistently associated with outcomes for neurocognitive endpoints but only occasionally with outcomes on motor tasks. Measures of later home environment, such as the family resource scale and the Henderson early learning process scale had limited effect, as would be expected for normally developing children at this age, and are not included.

We have focused on prenatal exposures, with postnatal hair concentration of MeHg a covariate in the analysis for prenatal effects. In a few tests this analysis suggested an adverse association with postnatal exposure in females. Since postnatal exposure differs substantially from prenatal exposure and since males are thought to be more susceptible, the interpretation of these findings is unclear. Analyses for the entire postnatal period are still in progress.

19 regression analyses revealed between one and five outlier scores involving a total of 40 different participants. In all cases, the association between prenatal MeHg exposure and the endpoint was the same, irrespective of whether outliers were included. For the two endpoints with a significant prenatal MeHg effect (Connor's teacher rating scale [two outliers] and the grooved pegboard non-dominant hand [three outliers]), the outlier scores all had prenatal MeHg concentrations of 7·5 μg/g or less and low performance.

Every model had between 0 and 3 influential points, defined as a score that may have affected the slope of the regression line but did not reach the status of an outlier. We report results with influential points included. However, the models with both interactions were re-run without influential points included, and in no case did the results for prenatal exposure change.

**Discussion**

Two of 21 endpoints were associated with prenatal MeHg exposure and developmental outcomes at 9 years of age. One association involved diminished performance (grooved pegboard non-dominant hand in males only) and the other an enhancement (hyperactivity index of the Connors teacher rating scale). As indicated by the distribution of p values in figure 3, both these outcomes are probably due to chance.

Results of studies of prenatal exposure to MeHg from seafood consumption in the Faeroe Islands\(^{22}\) and New Zealand\(^ {22,23}\) have shown adverse neuropsychological outcomes in school-aged children. The difference in findings from these studies and the Seychelles study has been investigated in two reviews\(^ {24}\) and several explanations have been proposed,\(^ {24}\) including the power of the studies to detect subtle differences. The assessment of standard power curves for the two studies shows that the power of the Seychelles study was only slightly less than that of the Faeroe Islands and both are substantially greater than any previous study. Our original power calculations estimated a 90% chance of detecting a five-point difference on the Bayley scales of infant development with every 10 μg/g increase in MeHg.

These outcomes might differ because of the cellular effect of very different concentrations of MeHg in the seafood consumed by these populations. The presence of cellular mechanisms in mammals that detoxify Hg raises the possibility that a larger bolus dose with a meal might behave differently than a small dose.\(^ {25}\) In Seychelles, the seafood consumed has a lower concentrations of MeHg than in the other two populations. The mean concentration of organic Hg in whale meat (the main source of MeHg in the Faeroe Islands) was 1−6 μg/g (SD 0−4).\(^ {26}\) In New Zealand the shark muscle consumed in the popular take-out food of fish and chips had a mean Hg concentration of 2·2 μg/g with some samples more than 4 μg/g.\(^ {27}\) By contrast, the Seychellois consume many different species of ocean fish, the mean MeHg content of which averages 0·3 μg/g (with 97·5% of the samples below 0·7 μg/g).\(^ {28}\)

Cord blood was used as the monitoring medium in the Faeroe Islands and the investigators argued that it is a more sensitive biomarker for prenatal MeHg exposure than concentrations in hair.\(^ {21}\) However, maternal hair has been the biological monitor of choice in most studies of prenatal exposure and was used in the New Zealand study.\(^ {29}\) Moreover, hair and blood concentrations are closely correlated, and hair can recapitulate exposure during the entire period of pregnancy.\(^ {12}\)

The tests used and the age at testing also differed between the studies.\(^ {27}\) However, our test battery included both global and domain-specific items and nearly all the tests reported previously had shown an association with MeHg.\(^ {22,23}\) Moreover, most of the tests gave results that were normally distributed and scores that were similar to those in Western countries, and were sensitive enough to detect the expected effects of covariates. The tests should have been sensitive enough to detect MeHg effects if they were present. A difference in age at testing is no longer a viable explanation since our previous findings at 66 months of age are now extended to 9 years, thus bracketing the 6 and 7 years ages used in the other studies.\(^ {6}\) One factor unique to the Faeroe Islands study is the consumption of whale meat and blubber. Whale blubber has high concentrations of polychlorinated biphenyls and other persistent organic pollutants and the meat has concentrations of inorganic Hg similar to MeHg.\(^ {26,29}\) These factors would not explain the associations reported from New Zealand.

Our study has potential limitations. As with any observational study, enrolment could have been biased. We sought to prevent such bias by offering participation to all children who reached the target age, and comparisons of those enrolled with those who did not participate did not show any bias. The best biomarker for prenatal exposure to MeHg is also controversial. In studies of neuro-developmental outcomes, choosing the appropriate
could serve as a sentinel population for fish consumers. Different results might have been obtained with different developmental tests, but we designed our tests to assess developmental domains known to be affected by prenatal MeHg poisoning and included most tests administered in previous studies. Finally, the choice of a different statistical approach might have led to different results. However, secondary analyses of our earlier studies with other statistical models have been consistent with conclusions based on the primary analysis. 3,10,31

In summary, the Seychelles Child Development Study longitudinal assessments at 9 years of age indicate no detectable adverse effects in a population consuming large quantities of a wide variety of ocean fish. These results are consistent with our earlier findings in the same children when examined at 6, 19, 29, and 66 months of age. In Seychelles, fetal exposure was continuous through frequent consumption of ocean fish containing concentrations of MeHg comparable to those consumed by the general population in the USA. We recorded effects from covariates known to affect child development, but did not find an association with prenatal mercury. We believe this finding is relevant to public health measures and that the Seychelles could serve as a sentinel population for fish consumers.

Contributors
G Myers, P Davidson, C Cox, T Clarkson, C Shamlaye, and D Palumbo designed and helped implement the study. C Shamlaye and G Myers managed the study in Seychelles and Rochester, respectively. P Davidson and D Palumbo trained testers and did site reliabilities. E Cernichiari analysed mercury and managed the data office. J Sloane-Reeves scored tests, trained testers, and with E Cernichiari managed the database. C Cox, G Widing, J Kost, and L-S Huang were responsible for statistical analysis. G Myers, P Davidson, C Cox, and T Clarkson interpreted the data and wrote the report.

Conflict of interest statement
None declared.

Acknowledgments
This research was supported by Grants R01-ES10219; R01-08442; ES-01247, and T32-ES-007271 from the US National Institutes of Health, the Food and Drug Administration, USDHHS, and by the Ministry of Health, Republic of Seychelles.

References


