

Consequences of Exposure Measurement Error for Confounder Identification in Environmental Epidemiology

Short title: Measurement error and confounding

Esben Budtz-Jørgensen^{1,2,*}, Niels Keiding¹, Philippe Grandjean^{2,3},
Pal Weihe^{2,4} and Roberta F. White^{2,3}

¹*Department of Biostatistics, University of Copenhagen
Blegdamsvej 3, DK-2200 Copenhagen N, Denmark.*

²*Institute of Public Health, University of Southern Denmark
Winslowparken 17, DK-5000 Odense C, Denmark.*

³*Departments of Environmental Health and Neurology,
Boston University Schools of Medicine and Public Health, Boston, MA 02118, USA*

⁴*Faroese Hospital System, FR-100 Tórshavn, Faroe Islands*

Editorial correspondence to:

*Esben Budtz-Jørgensen, Department of Biostatistics, University of Copenhagen,
Blegdamsvej 3, DK-2200 Copenhagen N, Denmark.

Phone: +45-35327927. Fax: +45-35327907. E-mail: ebj@biostat.ku.dk

SUMMARY

Non-differential measurement error in the exposure variable is known to attenuate the dose-response relationship. The amount of attenuation introduced in a given situation is not only a function of the precision of the exposure measurement but also depends on the conditional variance of the true exposure given the other independent variables. In addition, confounder effects may also be affected by the exposure measurement error. These difficulties in statistical model development are illustrated by examples from a epidemiological study performed in the Faroe Islands to investigate the adverse health effects of prenatal mercury exposure.

KEY WORDS: Bias, Confounding, Factor Analysis, Instrumental Variables, Measurement Error, Multiple Regression.

1. INTRODUCTION

In epidemiological studies, measurement error constitutes a serious problem. A uniform error on the exposure variable will simply bias the results toward the null hypothesis, and sensitivity analyses may be conducted to estimate the extent of this bias. However, oftentimes insufficient information on the measurement error is available to conduct proper sensitivity analyses. In addition, the effects of covariates may be of concern, and the confounder adjustment may also be affected by the measurement error.

A possible solution to these problems may be available when more than one exposure parameter is available, even if the measurement error for each of them is unknown. For instance, it may be possible to identify error variances by modeling the joint distribution of the exposure variables or some exposures may enter the analysis as instrumental variables. We illustrate these approaches using a data set from a study of neurobehavioral effects of prenatal exposure to methylmercury in a birth cohort of 1000 children.

2. THE FAROESE MERCURY STUDY

Methylmercury is a common contaminant in seafood and freshwater fish. While adverse effects have been unequivocally demonstrated in poisoning incidents, the health implications of lower-level exposures in fish eating populations have been controversial [1].

A birth cohort of 1022 children was generated in the Faroe Islands during 1986-1987 and is being studied prospectively to examine the possible adverse effects of prenatal exposure to methylmercury. The Faroese population is exposed to methylmercury

mainly through consumption of contaminated pilot whale meat. Information about the children's prenatal exposure was obtained by measuring mercury concentrations in maternal hair and cord blood at parturition. Both these exposure measurements are subject to short term biological fluctuations and to measurement error in the laboratory. However, because the latter biomarker was thought to be the best indicator of the causative concentration of mercury in the fetal circulation the analysis performed here focuses mainly on this exposure indicator.

Just before school entry the children underwent a detailed neuropsychological examination consisting of neuropsychological tests assessing different domains of brain function. Here we shall use as an example the California Verbal Learning Test (CVLT), a cognitive task which measures learning and memory. A list of 12 words that can be semantically clustered into four categories was given to the child over 5 learning trials. An interference list was then presented and finally the child was asked to recall as many of the words of the initial list as possible. The score on the test is number of correct answers, so the anticipated adverse effect of mercury would therefore decrease the score. Although the score may take only 13 different values, the distribution of this outcome variable, given important predictors, is well approximated by a normal distribution in the 900 children examined.

A possible reason why the unadjusted exposure-response relationship may be confounded in this study is that people living in the larger Faroese towns tend not to eat as much pilot whale meat as people living in the smaller fishing villages where whale meat may be more easily available. Grandjean et al. [2] controlled for a set of potential confounders that included sex and age of the child, maternal Raven score (a measure of intelligence) and socio-economic variables. Despite the inclusion of several relevant covariates, some degree of residual confounding is difficult to rule out. The study setting in the Faroe Islands is rather homogeneous, and the three main towns (Torshavn, Klaksvik or Tværá), where half of the 45,000 inhabitants live, would seem to differ little from smaller fishing villages nearby. Nonetheless, the three larger towns have easier access to day care, education, and employment, and a 'town' variable was therefore considered following discussion of possible residual confounding at an NIH workshop [3]. In routine multiple regressions (not taking exposure measurement error into account), town residents appeared to have an advantage for some outcomes [1]. However, as will be explained below, this may at least in part constitute an artifact caused by exposure measurement error and high correlation between exposure and the potential confounder.

3. BIAS CAUSED BY EXPOSURE MEASUREMENT ERROR

This section gives a brief review of the consequences of ignoring covariate measurement error in multiple regression analysis [4, 5]. In epidemiology the aim of the analysis is often to estimate the effect of the exposure (X) on the response (Y) for

fixed values of the column of confounders (Z). For that purpose multiple linear regression analysis has become a standard statistical tool. Thus, it is assumed that Y depends linearly on X and Z , i.e., $Y = \beta_0 + \beta_X X + \beta_Z^t Z + \epsilon$, where the residual error ϵ is distributed $N(0, \sigma_\epsilon^2)$. The exposure regression coefficient β_X is the parameter of main interest. The true exposure is often some sort of long-term average load which is impossible to measure precisely. Here attention is restricted to situations where X is observed indirectly through the error prone exposure measurement W , while the confounders Z are assumed to be directly observable. The relation between X and W is given by *the classical error model*, which assumes that W is given as a sum: $W = X + U$, where the measurement error U is independent of (X, Z) and normally distributed with mean zero and variance σ_U^2 . The measurement error is said to be *non-differential* if W and Y are independent given X and Z . If it is further assumed that the conditional distribution of X given Z is $N(\alpha_0 + \alpha_Z^t Z, \sigma_{X|Z}^2)$, then it is straightforward to show that

$$E(Y|W, Z) = c + \kappa\beta_X W + (\beta_Z + \beta_X(1 - \kappa)\alpha_Z)^t Z, \quad (1)$$

where $\kappa = \sigma_{X|Z}^2 / (\sigma_{X|Z}^2 + \sigma_U^2)$ and c is a (constant) intercept. Thus, if X is naively replaced by W in the regression analysis, then the estimated effects of both the exposure and the confounder are biased. The exposure effect is attenuated since clearly $0 < \kappa \leq 1$. The larger the measurement error variance (σ_U^2), the worse the attenuation. However, the attenuation also depends on the exposure-confounder relationship. If the association between these variables is strong, then the variance $\sigma_{X|Z}^2$ is small and the under-estimation of the exposure effect will be severe. On the other hand the attenuation is minimal ($\kappa = \sigma_X^2 / (\sigma_X^2 + \sigma_U^2)$) in designed studies where X and Z are independent. Because, the attenuated coefficient ($\kappa\beta_X$) is zero when the true coefficient (β_X) is zero, the naive test of no exposure effect is valid (the type I error rate is correct) but the test may not be efficient. Further, it is seen from (1) that naive estimates of confounder effects can be biased in either direction and that naive inference concerning these effects is generally invalid.

4. ADJUSTMENT FOR “TOWN” IN THE FAROESE DATA

The results of the previous section show that in studies where the exposure variable may be affected by measurement error, care is needed in the process of confounder selection.

In the Faroese mercury study the potential confounder *town* is a problematic covariate. Table I shows the estimated mercury effect on a child’s CVLT score using the log transformed cord blood mercury concentrations ($\log_{10}(B-Hg)$) as the exposure indicator. It is seen that when *town* is included in the set of confounders, the mercury effect is attenuated and the p -value changes from about 2% to about 6%. Further,

in accordance with the concern that the original set of confounders may not fully correct for Faroese social-class differences, it is estimated that urban children tend to do better than rural children on CVLT test ($\hat{\beta}_{town} = 0.364$, $p=0.060$). These results indicate that *town* may be an important confounder of the relationship between prenatal mercury exposure and childhood cognitive ability.

On the other hand, the cord blood mercury concentration does not give an error free reflection of the causative concentration of mercury in the fetal circulation or the fetal brain. As is often the case in environmental epidemiology, the exposure indicator is subject to ordinary laboratory measurement error as well as to biological fluctuations. Furthermore, because the village residents consumed more fish and whale meat, *town* must be strongly (negatively) related to the true prenatal mercury exposure. Regressing $\log_{10}(B-Hg)$ on the full set of confounders, it was estimated that the mercury exposure of an urban child is about 31% lower than for a rural child with the same values on the other confounders ($\hat{\alpha}_{town} = -0.164$). This exposure difference is highly significant, with a *t*-value of -6.1 . Had the true exposures been available, it is likely that an even stronger association would have been observed. From (1) it is seen that a consequence of the strong exposure difference between urban and rural children is that the attenuation of the mercury coefficient associated with adjustment for *town* could be due solely to exposure measurement error and not to a real effect of *town* on the response. Thus, even under the hypothesis of no *town* effect ($\beta_{town}=0$), a positive effect estimate of this variable would be expected because mercury exposure is negatively related both to *town* ($\alpha_{town} < 0$) and to the response ($\beta_{B-Hg} < 0$). In the following sections, two different statistical analyses are performed to take into account the mercury exposure measurement error providing a better basis for determining whether or not *town* is a confounder.

Table I: Estimated effects of a 10 fold increase in mercury exposure before and after correction for the effect of *town*. These estimates are obtained in a naive analysis ignoring measurement errors in the cord blood mercury concentration.

Town excluded		Town included	
$\hat{\beta}_{B-Hg}$	<i>p</i>	$\hat{\beta}_{B-Hg}$	<i>p</i>
-0.566	0.021	-0.466	0.063

5. METHOD OF MOMENTS

In studies where the exposure measurement error variance (σ_U^2) is known, it is straightforward to correct for the measurement error using the so-called *method of moments* [4, chapter 2.2]. In Section 3 it was shown that as the number of observations increases the naive least squares estimator ($\hat{\beta}_X$) converges to $\kappa\beta_X$, where

$\kappa = \sigma_{X|Z}^2 / (\sigma_{X|Z}^2 + \sigma_U^2)$. The denominator in the expression for κ is recognized as the conditional variance of W given Z ($\sigma_{W|Z}^2$), which can be estimated through a multiple regression analysis. Thus, a consistent estimate of β_X is obtained as $\hat{\beta}_X \cdot \{\hat{\sigma}_{W|Z}^2 / (\hat{\sigma}_{W|Z}^2 - \sigma_U^2)\}$. Fuller [4] provides an estimate of the variance of the method of moments estimator as well as modifications to the estimator when sample sizes are small.

In this section the cord blood mercury measurement error variance is estimated in factor analysis models whereupon the regression coefficients for the CVLT test are corrected using the method of moments.

In the Faroese study neither validation data (where the true exposure is measured directly) nor replication data (containing cord blood mercury measurements at different time-points) are available. Routine laboratory quality requires duplicate analysis of all samples and repetitive measurement of reference materials with known mercury concentrations. Thus, information is available on the analytical variation. However, since it is not reasonable to assume that the biological component of the measurement error is negligible, this information only provides a lower limit for the total error and is of minor value for the purposes of this paper. We therefore develop an alternative approach.

The main exposure variable used is the mercury concentration in cord blood obtained at the time of birth. Because methylmercury is thought to have a biological half-life of 45 days or slightly more, the concentration present in the cord blood reflects the exposure mainly during the last couple of months of gestation. A second exposure variable was also available, the mercury concentration in maternal hair (*H-Hg*). Hair grows approximately 1 cm per month, although this rate may be somewhat variable. Most of the hair samples were relatively long, and the 9-cm segment closest to the root was therefore analyzed to reflect the whole duration of the pregnancy. However, it may take 4-6 weeks for metal concentrations to show up in the hair strands beyond the scalp. This lag-time approaches the half-life of methylmercury in the body, and this exposure biomarker therefore mainly reflects the exposure during the period before the one that was the main determinant of the cord blood concentration. While hair color, hair growth rate and external contamination can affect mercury concentrations in hair, the two exposure biomarkers would be expected to correlate well if each individual's exposure was constant.

Under certain assumptions the measurement errors of the biomarkers can be identified by modeling their joint distribution. After a logarithmic transformation the relationship between the exposure variables is approximately linear (Figure 1). This leads to the following model for the exposure values of child i

$$\begin{aligned}\log_{10}(B-Hg)_i &= X_i + U_{B-Hg,i} \\ \log_{10}(H-Hg)_i &= \nu_{H-Hg} + \lambda_{H-Hg} \cdot X_i + U_{H-Hg,i}\end{aligned}$$

The model assumes that, except for measurement error, the indicators are given as linear functions of the true exposure X_i . The factor loading λ_{H-Hg} is included to avoid assuming that mercury exposures in blood and hair have the same scale. The intercept ν_{H-Hg} indicates the logarithm of the true hair concentration for a child with a true cord blood concentration of 1 unit. Furthermore, it is assumed that the measurement errors $U_{B-Hg,i}$ and $U_{H-Hg,i}$ are independent of each other and normally distributed with means 0 and variances σ_{B-Hg}^2 and σ_{H-Hg}^2 , respectively. The assumption of independence between measurement errors in cord blood and in maternal hair may (for instance) be appropriate if the true exposure is some sort of a long-term average mercury concentration. In this case the lag-time associated with the presence of mercury in the hair may ensure that the two biomarkers are not affected by the same random biological fluctuations on a temporal scale. In addition, concentrations of mercury in hair and in cord blood were determined by two different laboratories [6] which means that analytical errors are unlikely to be correlated.

If no distributional assumptions are made about X_i the model above is known as *the linear functional model*. However, in this section only *structural* models assuming that X_i has a normal distribution are considered. Including the unknown mean and variance of the true exposures six parameters have been introduced and therefore the model cannot be identified from the empirical means and covariances of the two biomarkers. In situations where the parameter of main interest is the factor loading (λ_{H-Hg}), orthogonal regression analysis is often applied assuming that the ratio between the measurement error variances is known. However, in this study no a priori information is available about this ratio. Furthermore, it does not seem attractive to make strong assumptions about the parameter of main interest in the estimation. A very simple model is obtained with $\lambda_{H-Hg} = 1$. This model assumes that the (median) ratio between measured mercury concentrations in maternal hair and cord blood is the same ($10^{\nu_{H-Hg}}$) for all true exposure levels. A priori this possibility may seem reasonable. However, because it will later be shown that the data contain important evidence *against* this hypothesis this model is not considered further.

In models where the parameter λ_{H-Hg} is different from 1, the measurement error variances of the two biomarkers are not directly comparable because they are on different scales. The measurement error standard deviation of the maternal hair concentration is converted to the scale of the cord blood concentration after division by the absolute value of the factor loading ($\sigma_{H-Hg}/|\lambda_{H-Hg}|$). Because the cord blood concentration is expected to give a more precise reflection of the true fetal exposure, the ratio between the converted measurement error variances of the blood and the hair concentration ($a = \lambda_{H-Hg}^2 \sigma_{B-Hg}^2 / \sigma_{H-Hg}^2$) is expected to be below 1. For fixed values of

the ratio of converted error variances Table II gives the estimated error variance of the cord blood concentration and the corrected mercury regression coefficient on the CVLT scores.

Table II: For fixed values of the ratio between converted error variances the table gives the estimated measurement error variance of the cord blood indicator and the corrected mercury regression coefficient on the CVLT scores

a	0.10	0.25	0.5	1	2	4	10
$\widehat{\sigma}_{B-Hg}^2$	0.007	0.014	0.022	0.030	0.038	0.044	0.049
$\widehat{\beta}_{B-Hg}$	-0.599	-0.636	-0.679	-0.734	-0.794	-0.848	-0.901

The results of Table II suggest that further modeling is needed. The value of the measurement error corrected mercury coefficient is too sensitive to the assumptions about the ratio of converted error variances. When this ratio is increased from 1/10 to 10, the estimated cord blood error variance increases by a factor of 6.8, which causes the corrected mercury coefficient to increase numerically by a factor of 1.5. In the next section more information is included and better estimates of σ_{B-Hg}^2 are obtained.

5.1 Inclusion of pilot whale consumption

The analysis of the joint distribution of mercury concentrations in cord blood and maternal hair showed that more information is needed in order to obtain good estimates of σ_{B-Hg}^2 . Additional information on the child's prenatal exposure to mercury is available from questionnaire data on maternal seafood intake during pregnancy. In connection with each birth, a midwife asked the mother about the average number of pilot whale dinners per month (*Whale*). After a transformation ($t(x) = \log_{10}(x+1)$) this variable is approximately linearly associated with the two other exposure indicators. Thus, the following factor analysis model is proposed

$$\begin{aligned} \log_{10}(B-Hg)_i &= X_i + U_{B-Hg,i} \\ \log_{10}(H-Hg)_i &= \nu_{H-Hg} + \lambda_{H-Hg} \cdot X_i + U_{H-Hg,i} \\ \log_{10}(Whale + 1)_i &= \nu_{Whale} + \lambda_{Whale} \cdot X_i + U_{Whale,i} \end{aligned}$$

where the measurement error terms are assumed to be mutually independent. With nine free parameters, the model is just identified from the observed means and covariances of the indicators. Closed form expressions for maximum likelihood estimators can be obtained for all parameters [4, chapter 1.5]. Table III shows estimated factor loadings and estimated measurement error variances in the proposed model for the exposure indicators.

The estimated measurement error variance of the cord blood mercury concentration is 0.018. Using the statistical software package Mx [7] a likelihood based 95%

Table III: Factor loadings, measurement error variances and converted error variances estimated in factor analytic model for indicators of prenatal mercury exposure.

Indicator	Loading (λ)	Error Variance (σ^2)	Converted Variance
$\log_{10}(B-Hg)$	1	0.018	0.018
$\log_{10}(H-Hg)$	0.840	0.036	0.052
$\log_{10}(Whale + 1)$	0.882	0.204	0.262

confidence interval was obtained for this error variance as (0.008; 0.027). From the estimated error variance, it is calculated that the coefficient of variation in the cord blood concentration is 30% ($\log_e(10)\sqrt{0.018}$). This result is approximately four times the documented analytical imprecision [6].

In the last column of Table III the measurement error variances have all been converted to the scale of the cord blood indicator. As expected, the concentration of mercury in the cord blood has the lowest converted variance. As indicated by the confidence interval given above, the converted error variance of the maternal hair concentration is significantly larger ($\chi_1^2=8.3$, $p=0.0039$). Thus, in this study the cord blood mercury concentrations is the best indicator of true prenatal mercury exposure. This result is in agreement with the observation of Grandjean et al. [8] that the cord blood concentration was a stronger predictor of childhood cognitive performance than the maternal hair concentration. Finally, we tested whether the ratio between mercury concentrations in cord blood and in maternal hair is constant ($\lambda_{H-Hg} = 1$). As announced above, this hypothesis was clearly rejected with a p -value below 0.0001.

Table IV: Estimated effects of a 10 fold increase in mercury exposure corrected for measurement error in the exposure using the method of moments with $\sigma_{B-Hg}^2=0.018$.

Town excluded		Town included	
$\hat{\beta}_{B-Hg}$	p	$\hat{\beta}_{B-Hg}$	p
-0.654	0.021	-0.542	0.063

The cord blood mercury regression coefficient was then corrected for the exposure measurement error using the method of moments assuming $\sigma_{B-Hg}^2=0.018$ (Table IV). It is seen that the method of moments estimator ($\hat{\beta}_X/\hat{\kappa}$) de-attenuates the naive mercury effect estimates, but because the estimation variance is increased by a factor of approximately $1/\kappa^2$ the p -values are unchanged. This result is typical after correction for non-differential exposure errors. Although effect estimates become stronger,

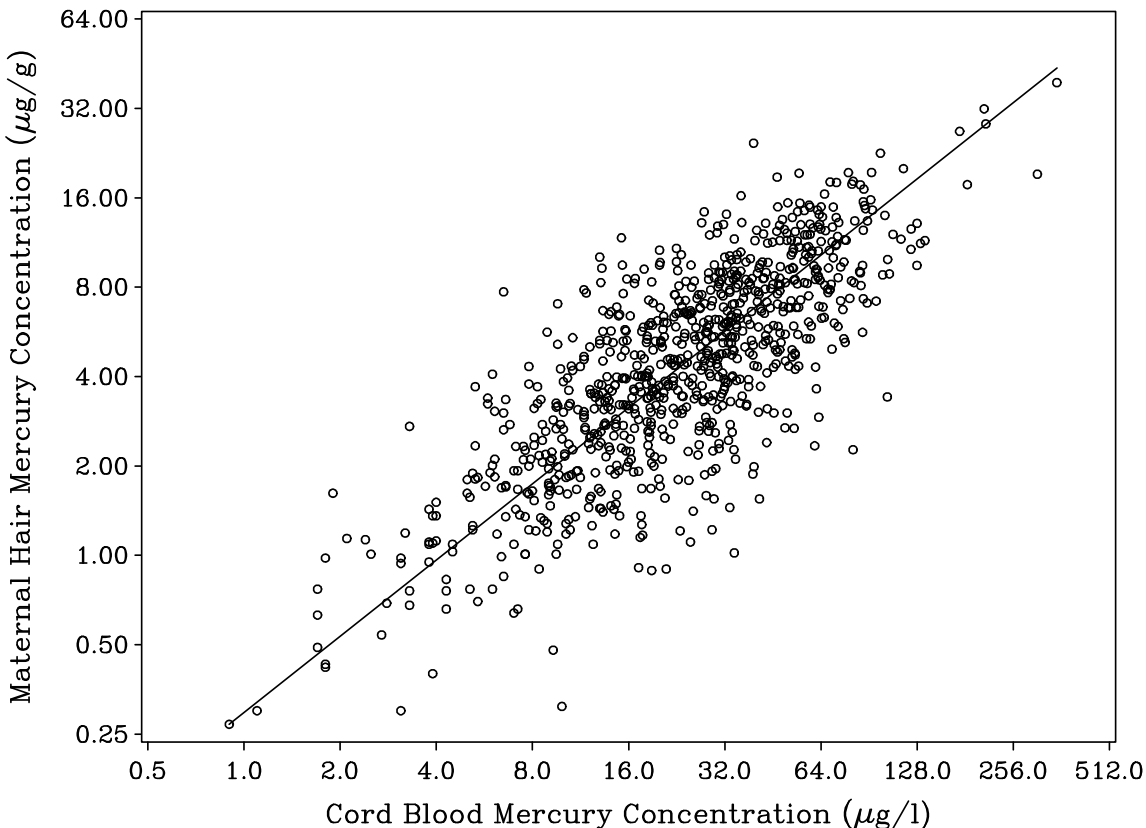


Figure 1: The relation between mercury concentrations in maternal hair and in cord blood estimated in the factor analysis model.

corrections often do not lead to important improvements in the statistical power to detect exposure effects. In this regard Tosteson and Tsiatis [9] showed that in generalized linear models the naive test of the hypothesis of no exposure effect is efficient if X is linearly related to W and Z .

Inclusion of the covariate *town* is seen to cause strong attenuation in the mercury coefficient, also after adjustment for exposure measurement error. Thus, *town* appears to be a real confounder which should be corrected for when assessing the effect of prenatal mercury exposure.

The above calculations are based on the (usual) assumption of independence between measurement error terms given the true exposure. However, some of the observed correlation between the two mercury biomarkers may have been caused by a mutual biological component in the error terms $U_{B-Hg,i}$ and $U_{H-Hg,i}$. (One such possibil-

ity would be a gross deviation in mercury load during the period (about 6 weeks pre-term) where the two biomarkers are both relevant. Such effects would lead to relatively modest correlations.) Unfortunately, the assumption of independent errors cannot be tested with the available data. Instead, the factor analysis was repeated assuming different values for the correlation between the two error terms. From these analyses, the measurement error variances were seen to increase as a function of the measurement error correlation (data not shown). For the cord blood concentration this increase was quite weak at low to moderate correlations, i.e., at a correlation of 0.2 the estimated error variance was 0.025. This variance is significantly smaller than the converted error variance of the maternal hair measurement. Thus, the conclusion that the cord blood concentration is the best indicator of prenatal mercury exposure is robust to the possibility of a moderate correlation between the error terms of the biomarkers.

The sensitivity of the decision on the role of the covariate *town* to the value of σ_{B-Hg}^2 was investigated by calculating how large this error variance would have to be for the *town*-attenuation in the mercury coefficient to be an artifact. Ignoring *town*, the naive regression coefficient of the mercury exposure was -0.566 (Table I), which according to the theory in Section 3 on average is equal to $\beta_X \cdot (\sigma_{W|Z}^2 - \sigma_{B-Hg}^2) / \sigma_{W|Z}^2$. Similarly, with adjustment for *town*, the estimated coefficient was -0.466 . If *town* is *not* a real confounder then the latter coefficient is on average given by the same expression as the former except that $\sigma_{W|Z}^2$ should be replaced by $\sigma_{W|Z,town}^2$ keeping β_X as before. Exploiting this, the expressions can be equated and solved for σ_{B-Hg}^2 . This yields an error variance of 0.104, which means that if the cord blood concentration is that imprecise then the observed *town*-attenuation in the mercury coefficient is a result simply of exposure measurement error. This variance is far above the upper confidence limit for σ_{B-Hg}^2 so it seems very unlikely that measurement error alone is responsible for the reduction in the estimated mercury effect observed when correcting for *town*.

6. INSTRUMENTAL VARIABLES

In the previous section the measurement error variance of the cord blood mercury concentration was estimated under the quite strong distributional assumptions of the factor analysis model. Here the effect of prenatal mercury exposure on the CVLT score is corrected for measurement error without direct estimation of σ_{B-Hg}^2 , thereby allowing weaker assumptions. This is done treating the two secondary exposure variables $\log_{10}(H-Hg)$ and $\log_{10}(Whale + 1)$ as so-called *instrumental variables*.

An instrumental variable T must satisfy three requirements

- I : T is related to X beyond what is explained by Z ($cov(X, T|Z) \neq 0$),
- II : T is independent of the measurement error on X ,

III : T is independent of Y given X and Z .

To see why these assumptions are sufficient for consistent estimation of β_X write $E(X|T, Z) = \gamma_0 + \gamma_T T + \gamma_Z^t Z$ and calculate the regression of Y on T and Z : $E(Y|T, Z) = E[E(Y|X, T, Z)|T, Z] = E[E(Y|X, Z)|T, Z] = \beta_0 + \beta_X E(X|T, Z) + \beta_Z^t Z = (\beta_0 + \beta_X \gamma_0) + \beta_X \gamma_T T + (\beta_X \gamma_Z + \beta_Z)^t Z$, where requirement III was used for the third expression. Thus, the regression coefficient of T is given as the product $\beta_X \gamma_T$, where the last factor (which is not zero by I) can be estimated consistently regressing W on T and Z since II yields $E(X|T, Z) = E(W|T, Z)$. An estimator of β_X is obtained by dividing the coefficient of T from the regression of Y with $\hat{\gamma}_T$. In the general case where more instrumental variables may be available Fuller [4, chapter 2.4] suggested the estimator:

$$\hat{\beta} = \tilde{\beta} \cdot c_1 + c_2 \quad (2)$$

where $\tilde{\beta}$ is a two-stage estimator obtained by first regressing W on Z and T and then regressing Y on Z and the expected values of W calculated from the first regression. The estimator $\tilde{\beta}$ is given as the coefficient of the expected exposures. The constants c_1 and c_2 are almost of no effect in the Faroese data ($c_1 \approx 1$, $c_2 \approx 0$), where n is large. Exact expressions for c_1 and c_2 can be derived from [4, pp. 151]. Fuller [4] showed that the estimator (2) is consistent under very weak assumptions on the distribution of $(\epsilon_i, U_{B-Hg,i}, X_i, W_i, Z_i)$.

Thus, in the instrumental variable analysis it is not necessary to assume independence between the measurement error $U_{B-Hg,i}$ and the residual error ϵ_i , nor is it necessary to assume a linear regression model for the association between the instrumental variables and the true mercury exposure. Furthermore, the two instrumental variables are allowed to be dependent given the true exposure but, as was the case in the factor analysis, maternal pilot whale consumption and the mercury concentration in maternal hair cannot contain information about the measurement error in the cord blood concentration.

Table V: Estimated effects of a 10 fold increase in mercury exposure corrected for measurement error in the exposure using instrumental variables.

Town excluded		Town included	
$\hat{\beta}_{B-Hg}$	p	$\hat{\beta}_{B-Hg}$	p
-0.673	0.031	-0.569	0.077

Table V shows estimated mercury effects corrected for measurement error using maternal hair mercury concentration and maternal pilot whale consumption as instrumental variables. Only one instrumental variable is needed, but inclusion of an

additional variable will lead to a more efficient estimate of the target parameter. It is seen that the results are in nice agreement with the method moments analysis, indicating that the stronger assumptions of the former analysis may be reasonable. Again, correction for *town* leads to an attenuated estimate of the exposure effect, reinforcing that this covariate should be included in the analysis.

7. DISCUSSION

In environmental epidemiological studies the effect of the exposure under investigation is typically quite weak. One way to increase power to detect such effects is to increase the variability in the exposure variable. Therefore, many studies consist of subjects from two different areas: an area of low-level exposed subjects and an area where the subjects are heavily exposed. The analysis presented here shows that this design is problematic especially if the exposure variable is not measured with excellent precision. Study power is only increased if it is not necessary to control for 'area', i.e., if subjects from different areas have comparable values on the response measure. This assumption can be checked by including 'area' in the set of confounders for control. In epidemiology the two most common methods for deciding whether adjustment for a potential confounder is necessary are the following. In one approach 'area' should be kept in the model if its effect on the response is significant. In another approach 'area' is considered an important confounder if its deletion is associated with a large change in the estimated exposure effect [10, chapter 15]. If the exposure variable is error prone both methods may lead to the wrong conclusion. Due to the measurement error and because 'area' (by design) is strongly correlated to the exposure, a positive effect of living in the low-level exposed area may be found even in the situation where the design is successful and study subjects are comparable across areas. Furthermore, adjustment for 'area' may lead to an important (numeric) decrease in the estimated exposure effect. Thus, no matter which method is used 'area' may wrongly be included in the analysis, which can lead to severe under-estimation of the exposure effect. In studies of such designs it is therefore particularly important to obtain data on the size of the measurement error in the exposure variable.

In the Faroese mercury study, correction for measurement error in the prime exposure variable was obtained by including two secondary exposure indicators in the analysis. This was done in two different ways and both analyses indicated that the covariate *town* is a real confounder of the association between prenatal mercury exposure and childhood cognitive ability. For the CVLT test, the mercury regression coefficient is virtually unchanged after adjustment for both measurement error and this additional confounder, but a small increase in the variance of the coefficient results in a change in *p*-value from about 2% to about 6%. In fact, the outcome variable used as an illustrative example in this study was chosen for this very reason. For other indicators of neurobehavioral deficit the mercury associations are rather robust [1], and the overall

mercury effect is therefore not in any way insignificant after adjustment for *town* [11].

ACKNOWLEDGEMENTS

This study was supported by grants from the National Institute of Environmental Health Sciences (ES06112 and ES09797), the U.S.Environmental Protection Agency (9W-0262-NAEX), the European Commission (Environment Research Programme), the Danish Medical Research Council (22-00-0540), and the Danish Health Insurance Foundation (2000B084). The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the NIEHS, NIH or any other funding agency.

REFERENCES

1. National Academy of Sciences (NAS). *Toxicological Effects of Methylmercury*. National Academy Press, 2000.
2. 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. *Neurotoxicology and Teratology* 1997; **19**: 417-428.
3. National Institutes of Health (NIH). *Scientific Issues Relevant to Assessment of Health Effects from Exposure to Methylmercury*. 1998.
4. Fuller WA. *Measurement Error Models*. Wiley, 1987.
5. Carroll RJ, Ruppert D, Stefanski LA. *Measurement Error in Nonlinear Models*. Chapman & Hall, 1995.
6. Grandjean P, Weihe P, Jørgensen PJ, Clarkson T, Cernichiari E, Viderø T. Impact of maternal seafood diet on fetal exposure to mercury, selenium, and lead. *Archives of Environmental Health* 1992; **47**: 185-195.
7. Neale MC, Boker SM, Xie G, Maes HH. *Mx: Statistical Modeling*. Box 126 MCV, Richmond, VA: Department of Psychiatry. 5th Edition, 1999.
8. Grandjean P, Budtz-Jørgensen E, White RF, Jørgensen PJ, Weihe P, Debes F, Keiding N. Methylmercury exposure biomarkers as indicators of neurotoxicity in children aged 7 years. *American Journal of Epidemiology* 1999; **150**: 301-305.
9. Tosteson T, Tsiatis A. The asymptotic relative efficiency of score tests in a generalized linear model with surrogate covariates. *Biometrika* 1988; **75**: 507-514.

10. Rothman KJ, Greenland S. *Modern Epidemiology*. Lippincott-Raven, 1998.
11. Budtz-Jørgensen E, Keiding N, Grandjean P, Weihe P. Estimation of health effects of prenatal mercury exposure using structural equation models. *Environmental Health* 2002; **1**: 2.