Underestimation of risk due to misclassification

4 authors, including:

Esben Budtz-Jørgensen
University of Copenhagen
122 PUBLICATIONS   4,948 CITATIONS

SEE PROFILE

Niels Keiding
University of Copenhagen
412 PUBLICATIONS   18,628 CITATIONS

SEE PROFILE

Pal Weihe
The Faroese Hospital System
176 PUBLICATIONS   9,695 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Epidemiological study of Human Reproduction View project

Project UNICLAS: University of Copenhagen Longitudinal Ageing Study View project

UNDERESTIMATION OF RISK DUE TO EXPOSURE MISCLASSIFICATION

PHILIPPE GRANDJEAN[1,2], ESBEN BUDTZ-JØRGENSEN[1,3], NIELS KEIDING[3] and PAL WEIHE[1,4]

[1] Institute of Public Health
University of Southern Denmark
Odense, Denmark

[2] Department of Environmental Health
Boston University School of Public Health
Boston, MA, USA

[3] Department of Biostatistics
Panum Institute
University of Copenhagen
Copenhagen, Denmark

[4] Department of Occupational and Public Health
Faroese Hospital System
Tórshavn, Faroe Islands

Abstract. Exposure misclassification constitutes a major obstacle when developing dose-response relationships for risk assessment. A non-differential error results in underestimation of the risk. If the degree of misclassification is known, adjustment may be achieved by sensitivity analysis. The purpose of this study was to examine the full magnitude of measurement error in determining the prenatal exposure to methylmercury. We used data from a prospective study of a Faroese birth cohort. Two biomarkers of methylmercury exposure were available, i.e., the mercury concentrations in cord blood and in maternal hair (sampled at the time of parturition). The laboratory imprecision on both chemical analyses was thought to be below 5% coefficient of variation (CV). As a third exposure parameter, we used the dietary questionnaire response on frequency of whale meat dinners. Factor analysis and structural equation analysis were applied to assess the full extent of the imprecision. The calculated total imprecision much exceeded the known laboratory variation: the CV was 28-30% for the cord-blood concentration and 52-55% for the maternal hair concentration. The dietary questionnaire response was even more imprecise. These findings illustrate that measurement error may be greatly underestimated if judged solely from reproducibility or laboratory quality data. Adjustment by sensitivity analysis is meaningful only if realistic measurement errors are applied. When exposure measurement errors are overlooked or underestimated, decisions based on the precautionary principle will not appropriately reflect the degree of precaution that was intended.

Key words:
Biological markers, Dose-response relationship, Prenatal exposure delayed effects, Risk factors, Statistical models

INTRODUCTION

In interpreting the Precautionary Principle, one of the potential stumbling points is how much uncertainty or doubt is acceptable [1]. This issue is further complicated when sources of uncertainty are overlooked or underestimated, so that the degree of uncertainty itself is uncertain. In regard to environmental health problems, exposure misclassification is a major obstacle to obtaining accurate dose-response relationships. Thus, the extent to which a true association can be observed depends on the precision with which the key parameters are assessed, in particular the exposure to the suspected causal agent. Uncertainty therefore has crucial implications regarding causal inference, where the strength of association is a key criterion [2].

Random or non-differential exposure measurement error results in underestimation of the risk. The larger the imprecision of the exposure parameter, the greater the underestimation. If the degree of imprecision is known, then adjustments can be made by sensitivity analysis. However, because a gold standard is rarely available, the degree of misclassification is seldom known for sure, and its magnitude must therefore be assumed, e.g., from comparisons between exposure indicators. An important question is whether such assumptions are reasonable.

The importance of exposure misclassification is widely recognized, especially in regard to regression analysis [3–5]. Research studies therefore aim at providing exposure parameters of high validity and with the highest possible precision. Exposure biomarkers have recently become a key instrument for this purpose [6]. An exposure biomarker is usually a xenobiotic compound – or a metabolite – within the body, but may also be an interactive product between the compound – or metabolite – and an endogenous component, or another event related to the exposure. Validation studies usually involve comparison of different methods, analysis of reference materials, and replication studies. The imprecision is generally expressed as the coefficient of variation (CV), i.e., the standard deviation divided by the mean of a large number of analyses of a uniform sample.

The validity of exposure parameters must also be considered in regard to other factors, such as the exposure-dose relationship and the degree to which an exposure parameter reflects the concentration at the sensitive target or compartment within the body (in addition, possible time-dependence of the exposure coupled with temporal variations in susceptibility to toxic effects). Thus, laboratory accuracy and precision may not reflect the total degree of uncertainty associated with a particular exposure parameter.

We have examined this issue in regard to methylmercury, an important environmental health hazard, where uncertainty appears to affect the risk evaluation [7]. Methylmercury is a global water contaminant that is accumulated in freshwater fish and seafood; the most commonly used exposure biomarkers are the mercury concentrations in scalp hair and in blood [8]. Toxic effects on brain development are thought to constitute the critical effect of this toxicant [7]. We have used data from a prospective study of a Faroese birth cohort to estimate the true extent of imprecision associated with different methylmercury exposure parameters.

MATERIALS AND METHODS

Cohort examinations

A birth cohort of 1022 children was established in the Faroe Islands in 1986–1987, and the prenatal exposure to methylmercury was estimated by measuring mercury concentrations in cord blood and in maternal hair sampled at the time of parturition [9]. The laboratory imprecision on both chemical analyses was thought to be below 5% CV, with individual results often in the 2–4% range. In addition, a high quality was secured by participation in external quality assessment schemes with highly satisfactory results. The two exposure biomarkers showed a skewed distribution and were therefore logarithmically transformed. The transformed results showed an approximately linear relationship with an excellent correlation (Fig. 1). In connection with childbirth, a dietary questionnaire was administered by the midwife. The number of whale meat dinners per month during pregnancy, the main source of methylmercury exposure in this community, was closely associated with both mercury biomarkers [9].

Methylmercury in whole blood mainly represents recent exposures, i.e., during the third trimester of pregnancy [8]. Hair samples were in most cases about 9 cm long, thus reflecting hair growth during the full pregnancy; the mercury concentration therefore represented the average maternal exposure during gestation [8]. In addition to temporal variations, external contamination and other factors may affect the mercury concentration and therefore contribute to the variability of the biomarkers [8,10]. However, the extent to which such factors play a role is unknown.

A total of 917 cohort children were subsequently examined by neurobehavioral methods at the age of 7 years; mercury-associated deficits were observed in several domains of brain function [11]. As an indicator of fetal methylmercury
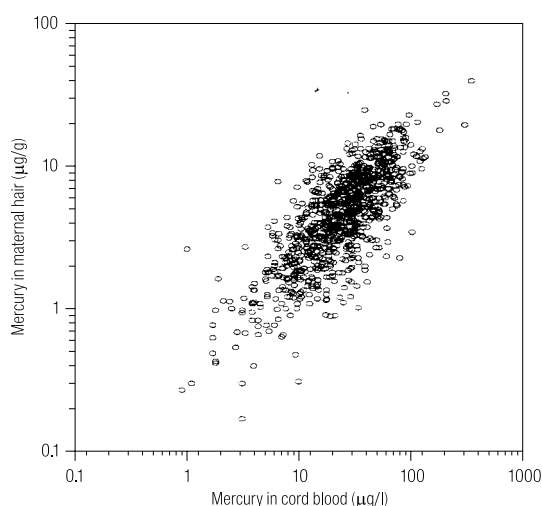
Fig. 1. Association between two biomarkers of prenatal exposure to methylmercury: mercury concentrations in cord blood and maternal hair sampled at the time of parturition [9]. The high correlation between the two parameters ($r = 0.78$) suggests that the validity of the biomarkers may be high. However, as no gold standard is available, the true imprecision of these measurements is unknown.

exposure, the umbilical cord-blood mercury concentration generally appeared to be the best risk predictor.

Statistical analyses

The true exposure is not known but is observed through the error-prone exposure variables. According to the classical error model the observed exposure is a sum of the true exposure and a random error with mean 0. This error is said to be non-differential if it is independent of the effect variable under consideration. Failure to account for non-differential error will generally lead to underestimation of exposure-related effects [3–5]. Unbiased effect estimation requires information about the size of the exposure imprecision.

Two approaches to assessing the variance of the error component in each of the mercury exposure variables were employed. The factor analysis model [4,12] assumes that (after logarithmic transformation) each exposure parameter Hg_i is linearly related to a common latent factor Hg representing the true methylmercury exposure level:

$$\log(Hg_i) = \alpha_i + \beta_i * \log(Hg) + \varepsilon_i$$

where:
α_i is an intercept (for exposure parameter i), and β_i is the loading factor that allows different scales for different ex-

posure parameters. In the current analysis, the cord blood concentration was chosen as the reference indicator. For this variable, parameters are therefore fixed at $\alpha_{blood} = 0$ and $\beta_{blood} = 1$. The restriction for β_{blood} means that the true mercury exposure is expressed on the same (log-transformed) scale as the cord blood concentration. The error terms ε_i are assumed to be independent and normally distributed with an average of 0.

The factor analysis model requires at least three exposure indicators for all unknown parameters to be identifiable [4,12]. In addition to the two prenatal biomarkers, i.e., the mercury concentrations in cord blood and in maternal hair, the dietary questionnaire response on the number of pilot whale meat dinners per month was included as the third variable. The latter parameter was also converted to a logarithmic scale (after adding 1 to the questionnaire response).

In models where factor loadings (β_i) are not equal, the measurement error variances (variance of ε_i) are not directly comparable, because they are on different scales. Thus, an indicator with a low error variance my be imprecise if the factor loading is also low.

Meaningful comparisons of precisions may be obtained by converting the measurement error of each variable to the scale of the cord blood concentration. This is achieved by multiplying the estimated error standard deviation by the inverse of the factor loading. Finally, converted imprecisions my also be expressed as CV values for the untransformed concentration. Because the logarithmic (base 10) transformation was used, the CV is the square root of the converted error variance multiplied by 2.30 (the natural logarithm of 10).

A related, but more sophisticated approach is structural equation analysis [13], which has recently become feasible with the availability of software packages [14,15]. With this method, study variables are first divided into covariates and response variables. The covariates are assumed to be error free while the response variables are allowed to be affected by measurement error. The aim is to model the conditional distribution of the response variables given the covariates. This is achieved by viewing the response variables as manifestations of a limited number of latent

variables and by assuming linear regressions between latent variables and observed covariates. Thus, the model consists of two parts.

The measurement model describes each of the latent variables, including the exposure variable. For this purpose response variables are grouped. Variables in the same group are assumed to be related to the same latent variable as in the factor analysis. The structural model characterizes the relation between the latent variables while taking into account effects of the covariates. Using the data from the Faroese birth cohort, this approach has been used to estimate the effects of prenatal methylmercury on overall neurobehavioral performance at age 7 years [15].

Again, the latent exposure variable is modeled from the two exposure biomarkers and the questionnaire information. The measurement model includes one group of neurobehavioral test scores that measure and underlying motor function, while another group are considered indicators of a latent verbally-mediated function [15]. Potential confounders of the relation between mercury exposure and neurobehavioral performance are assumed error free and included as covariates. Thus, contrary to the simple factor analysis this approach exploits information also from the confounders and the effect variables when estimating the true exposure levels.

RESULTS

The results of the factor analysis are shown in Table 1. Relative to the cord-blood parameter, the loading factor of the hair parameter is less than 1, but greater differences are seen in the measurement error variances. In the third column of the Table, both error variances are expressed on the scale of the cord-blood parameter. A likelihood-based 95% confidence interval for the error variance of the cord-blood parameters is 0.0078–0.0265. The converted error variance of the maternal hair concentration is significantly larger than the one for the blood parameter ($\chi^2 = 8.35$, df = 1, p = 0.0039). The questionnaire response is even more imprecise. Thus, the cord-blood mercury concentration is therefore the best, i.e., the least imprecise, indicator of true prenatal mercury exposure.

The results of the structural equation model (Table 2) are quite similar to those based on the simple factor analysis, and the cord-blood parameter again appears to be much less imprecise than the hair analysis. The structural equation results also take into account the effects of confounders and outcome variables. However, in comparison with the individual exposure parameters, these parameters are only weakly associated with the latent methylmercury exposure. Thus, these results are in good agreement with the factor analysis model.

Table 1. Results of factor analysis of three parameters that reflect the prenatal exposure to methylmercury [12]

Indicator	Loading factor	Error variance	Converted variance	Coefficient of variation
Blood mercury	1	0.018	0.018	0.30
Hair mercury	0.840	0.036	0.052	0.52

Each parameter has been log-transformed. Because factor loadings are not equal, the two mercury exposure biomarkers are on different scales, and the error variances are therefore not directly comparable. After expression on the scale of the cord-blood mercury concentration, the converted variance is then used to calculate the coefficients of variation (see text)

Table 2. Results of structural equation analysis of two biomarkers of prenatal methylmercury exposure (cf. Table 1), while taking into account also counfounders and outcome variables [15].

Indicator	Loading factor	Error variance	Converted variance	Coefficient of variation
Blood mercury	1	0.015	0.015	0.28
Hair mercury	0.809	0.038	0.058	0.55

Each exposure parameter has been log-transformed

The most important finding is the overall level of imprecision. Given the high analytical quality of the mercury measurements and their wide acceptance as valid biomarkers [8], a relatively small total error would be anticipated. However, the calculated CVs much exceed the known laboratory variations. The total error for the two mercury concentrations is at least five-fold greater than the laboratory imprecision, and the questionnaire response is even more imprecise.

DISCUSSION

The results of these analyses are in agreement with the observation that the cord-blood mercury concentration

is a stronger predictor of cognitive deficits than the maternal hair concentration [11]. However, the magnitude of estimated total imprecision is far in excess of previous determinations of laboratory variability. The imprecision of each exposure biomarker can be looked upon as a sum of two different types of error: laboratory measurement imprecision and biological or pre-analytical variation. The second error component arises because the mercury concentration in the fetal circulation is not constant over time but varies according to maternal mercury intake. It may also include individual differences in the distribution of mercury in the body and across the placenta, variations in the hematocrit of the fetal blood, etc. Hair concentrations may depend on hair structure and growth rates. Because hair is exposed to the external environment for several weeks or months before collection of the specimen, contamination from external mercury vapor or hair treatment products may add to the mercury concentration originally incorporated in the hair root [8]. The total magnitude of these factors had not been previously assessed.

These findings illustrate that measurement error may be greatly underestimated if judged solely from laboratory quality data. This finding is important, because a correct classification of the exposure is a key to obtaining appropriate information on the exposure-associated effects. Although sensitivity analyses may provide adjustment for the imprecision in the exposure assessment, such adjustments are only meaningful if they include measurement errors of a realistic magnitude.

If based on laboratory performance, an imprecision estimate is therefore obtained with regard to a parameter which does not truly reflect the factor that it purports to indicate. The precision is therefore confused with the quality of the information. In the present example, the mercury concentration in maternal hair or some other sample is erroneously assumed to represent the dose at the fetal target organ. Such problems are sometimes referred to as a Type III error, i.e., providing an answer to the wrong question [16].

Given the magnitude of the total imprecision observed in this study, limited benefit will arise from intensified efforts to improve the quality of the chemical analyses in the labo-

ratory. The analytical imprecision is already only a small part of the total error. Thus, the pre-analytical sources of imprecision deserve increased attention.

Regression analysis generally assumes that the exposure variable has been assessed without error. Any non-differential error will bias the results toward the null hypothesis. In the case of biomarker analyses with a low detection limit, the imprecision is usually random. Thus, regression coefficients and odds ratios will be attenuated. Additional bias toward the null may occur in the presence of confounders [12]. On the other hand, if the population examined represents a very wide exposure range, then the bias will be reduced.

Weak associations, e.g., a relative risk less than 2.0, are more likely to be dismissed as not being in favor of a particular causal relationship. In such cases, studies may reach completely opposite conclusions in regard to causality [17]. Methylmercury is an appropriate example in this regard, as current environmental and dietary exposure limits vary by a factor of about 5, although presumably based on the same evidence [7].

Environmental mercury toxicity should also be considered in a wider perspective of developmental neurotoxicity. Lead and polychlorinated biphenyls (PCBs) are two other environmental pollutants also known to cause these effects, but evidence is meagre on the role of other chemicals. Despite the suspected vulnerability of the fetal brain, about 150 chemicals have been documented as neurotoxic to adults [18]. Still, these compounds represent but a fraction of the chemical universe. It would therefore seem highly unlikely that science has already discovered all important developmental neurotoxicants. Thus, methylmercury is probably just one out of many chemicals that are toxic to the developing brain. The evidence on methylmercury, as well as the one on inorganic lead and PCBs, indicates that measurement error and other sources of uncertainty severely hamper the ability to draw straightforward conclusions.

When decisions on preventive efforts are made, they should be justified by the overall scientific database and take into account its various uncertainties and inconsistencies. A precautionary approach is warranted when

incomplete information is available and when uncertainty is a major consideration. If uncertainty is misjudged and thought to be smaller than its true magnitude, even a precautionary approach may be misguided. In this regard, a recent recommendation appropriately suggested that it is necessary to search out and address 'blind spots' and gaps in scientific knowledge [19]. This search must include an in-depth evaluation of the total extent of imprecision of exposure assessments applied for determining causal associations and dose-response relationships.

CONCLUSIONS

Although commonly used to represent the imprecision of exposure assessments, laboratory reproducibility may greatly underestimate the total imprecision of exposure biomarkers. In epidemiological studies, unrecognized random error will result in underestimation of the true effects of the exposure. Using data from a large birth cohort study, biomarkers of prenatal methylmercury exposure were found to be at least five-fold more imprecise than suggested by laboratory quality data. Factor analysis and structural equation modeling were found to be useful approaches to ascertain the full extent of the imprecision. Adjustment by sensitivity analysis is meaningful only if realistic measurement errors are applied. Because serious underestimation of exposure imprecision is likely to be common, the significance of environmental hazards may be misjudged. When exposure measurement errors are overlooked or underestimated, decisions based on the precautionary principle will not appropriately reflect the degree of precaution that was intended.

REFERENCES

1. Cameron J, Wade-Gery W. *Addressing uncertainty. Law, policy and the development of the precautionary principle.* In: Dente B, editor. *Environmental Policy in Search of New Instruments.* Dordrecht: Kluwer; 1995. p. 95–142.

2. Hill AB. *The environment and disease: association or causation?* Proc R Soc Med 1965; 8: 295–300.

3. Copeland KT, Checkoway H, McMichael AJ, Holbrook RH. *Bias due to misclassification in the estimation of relative risk*. Am J Epidemiol 1977; 105: 488–95.

4. Fuller WA. *Measurement Error Models.* New York: Wiley; 1987.

5. Carroll RJ, Ruppert D, Stefanski LA. *Measurement Error in Nonlinear Models*. New York: Chapman and Hall; 1995.

6. Grandjean P. *Biomarkers in epidemiology*. Clin Chem 1995; 41: 1800–3.

7. Grandjean P. *Mercury risks: controversy or just uncertainty?* Publ Health Rep 1999; 114: 512–5.

8. Grandjean P, Jørgensen PJ, Weihe P. *Validity of mercury exposure biomarkers*. In: Wilson SH, Suk WA, editors. *Biomarkers of Environmentally Associated Disease*. Boca Raton, FL: CRC Press/Lewis Publishers; 2002. p. 235–47.

9. 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.* Arch Environ Health 1992; 47: 185–95.

10. National Research Council. *Toxicological effects of methylmercury.* Washington, DC: National Academy Press; 2000.

11. Grandjean P, Weihe P, White RF, Debes F, Araki S, Murata K, et al. *Cognitive deficit in 7-year-old children with prenatal exposure to methylmercury.* Neurotoxicol Teratol 1997; 19: 417–28.

12. Budtz-Jørgensen E, Keiding N, Grandjean P, Weihe P, White RF. *Consequences of exposure measurement error for confounder identification in environmental epidemiology.* Stat Med 2003; 22(19): 3089–100

13. . Bollen KA. *Structural Equations with Latent Variables.* New York: Wiley; 1989.

14. Muthén LK, Muthén B. Mplus. *The Comprehensive Modeling Program for Applied Researchers. User's Guide*. Los Angeles: Muthén & Muthén; 1998.

15. Budtz-Jørgensen E, Keiding N, Grandjean P, Weihe P. *Estimation of health effects of prenatal mercury exposure using structural equation models.* Environ Health 2002; 1: 2.

16. Schwartz S, Carpenter KM. *The right answer for the wrong question: consequences of type III error for public health research*. Am J Public Health 1999; 89: 1175–80.

17. Weed DL. *Underdetermination and incommensurability in contemporary epidemiology*. Kennedy Inst Ethics J 1997; 7: 107–27.

18. Grandjean P, Sandoe SH, Kimbrough RD. N*onspecificity of clinical signs and symptoms caused by environmental chemicals*. Hum Exp Toxicol 1991; 10: 167–73.

19. European Environment Agency. *Late Lessons from Early Warnings: the Precautionary Principle 1896– 2000.* Environmental issue report No. 22, 2001.Copenhagen: European Environment Agency; 2001.