Interindividul Variability in Modeling Exposure and Toxicokinetics: A Case Study on Cadmium

Wout Slob and Ernõ I. Krajnc
National Institute of Public Health and Environmental Protection, 3720 BA Bilthoven, the Netherlands

Exposure assessments are often based on characteristics of the average individual (e.g., average body weight or average consumption of a certain food product). When toxicokinetic models are used to estimate internal doses from external doses, the parameters of the model relate to the average individual as well. Such calculations therefore result in predictive statements that relate to the average individual in the population. However, protection of the average (median) individual implies protection of only 50% of the population. In view of the large variation that typically exists in human populations with respect to behavioral and biological properties, risk calculations based on the average individual may lead to quite misleading results. Nonaverage individuals should therefore be evaluated as well. This paper illustrates by a case study on cadmium how sources of interindividual variation may be quantified and how this information can be used in a model-based risk analysis.

Ingestion by food is a primary source of human exposure to many chemicals (1). To describe interindividual variability in intake caused by differences in consumption habits, a statistical exposure model (STEM) has been developed (2). This model succinctly describes the lifelong intake of persistent chemicals for a population as a whole:

\[ Y_i(t) = F(t) \eta_i \]

where \( Y_i(t) \) denotes individual \( i \)'s daily intake of the chemical at age \( t \), \( F(t) \) is the daily intake by the median individual at age \( t \), and the term \( \eta_i \) represents the individual's deviation from the typical consumption pattern \( F(t) \); \( \eta_i \) is assumed to be lognormally distributed. Thus, the intake of a chemical by the population as a whole is described by a lognormally distributed bundle of functions, each function relating to the intake over a lifetime of a single individual. Note that in this representation intake of cadmium from food is described continuously rather than intermittently at meals. This approximation is valid because of the time-scale considered here (viz, the human life span).

The parameters of the function \( F(t) \) and the coefficient of variation of \( \eta_i \) can be derived from food consumption surveys that have been carried out in several countries. For example, in the Netherlands such a survey (3,4) resulted in detailed consumption data recorded on 2 consecutive days from a representative sample of 6000 persons (response rate 80%). These data have been used to investigate the assumptions of STEM (2). The assumptions were largely fulfilled, e.g., the distribution of log-intakes appeared to be quite well described by a normal distribution, whereas their variance was nearly homogeneous over age. In cases where the assumptions were not completely fulfilled, this appeared to have no major consequences, as shown by simulation studies (2). The representation given by Equation 1 was therefore maintained for its simplicity. Figure 1 shows the model applied to intake of cadmium in the Dutch population; the intake of the whole population in relation to age is illustrated. Furthermore, the fraction of the population exceeding any critical intake value can be easily derived from the model. In the case of cadmium, a provisional tolerable weekly intake (PTWI) has been formulated at 400–500 µg per week or 7 µg per kg body weight (5). Figure 1 shows that even the high percentiles of the intake distribution are below these limits. In this study STEM serves as the input for a toxicokinetic model to estimate the frequency distribution of cadmium concentrations in the kidney (the primary target organ). Inhalation of cadmium from smoking is also considered.

We used an existing toxicokinetic model for cadmium developed by Nordberg and Kjellström (6). This classical compartmental model consists of eight compartments showing first-order kinetics. Note that in contrast to physiologically based models, the parameters have no measurable physiological correlates; their values were determined by calibration to Swedish and Japanese data of cadmium concentrations in various compartments, measured in the general population and in specific exposure groups (e.g., smokers, workers). Yet, some physiological knowledge is incorporated in the model. For example, the blood is represented by three compartments: one for binding to metallothionein, one for binding to other proteins (albumin), and one for accumulation in erythrocytes. The fraction of cadmium binding to metallothionein is assumed to have a maximum, resulting in nonlinearity of the model. Also, some of the transfer rate coefficients are time dependent (e.g., the reabsorption rate of the kidney decreases at higher ages). The model does not include the induction of metallothionein; it is not clear if this mechanism plays an important role at current exposures in the human population. Nonetheless,
the model fits the data for various exposure levels (Swedish/Japanese/Dutch populations; smoking/nonsmoking; work environments). The model, with the same parameter values as assessed by Nordberg and Kjellström (i.e., based on Swedish and Japanese data) appeared to fit the available Dutch data (7) quite well. Model evaluations (i.e., the simultaneous integration of the differential equations) were done using the software package SimuSolv (Dow Chemical Company, Midland, Michigan), using Gear's integration algorithm.

Obviously, interindividual variation in intakes will lead to interindividual variation in internal doses. The model formulation of STEM enables one to evaluate the consequences of variation in intake for the internal concentrations by a Monte Carlo analysis using the following procedure: 1) randomly select a value for η (Equation 1) from its lognormal distribution, 2) multiply this value of η with the median exposure function Y(t) resulting in a function of age, Y(η), representing the intake over a lifetime of a randomly selected individual from the population, 3) present the function Y(η) as an input to the toxicokinetic model and evaluate that individual's cadmium concentration in the compartment of interest (e.g., the kidney), 4) repeat steps 1–3 a large number of times (in this study 500 or 1000), and 5) evaluate the variation in the model outcomes [in terms of the coefficient of variation (CV)].

The variation in cadmium (absolute) intake, estimated by STEM at CV = 22%, resulted in concentrations in the kidney having a CV of 23%, at ages where the maximum kidney concentration is predicted. This is considerably less than the variation in kidney concentrations observed at comparable ages (CV = 83%) (7). Apparently, sources of variation other than intake play a substantial role.

The state variables in the cadmium model of Nordberg and Kjellström are formulated in terms of amounts. To arrive at concentrations these amounts must be divided by the relevant organ weights. Variation in organ weights thus introduces another source of variation between individuals that is relevant here. To quantify this variation, the following approach was adopted. Body weights in the population were described by a statistical model similar to Equation 1, resulting in a CV of 17% for interindividual variation. This variation was assumed to hold for the variation in organ weights as well. The Monte Carlo analysis was then extended by randomly selecting a kidney weight function (of age) in each run, which was used to translate predicted amounts into concentrations. Together with the variation in intakes, this resulted in a CV for predicted concentrations in the kidney of 28%, not much higher than the previous result.

The parameters of the toxicokinetic model, reflecting an individual's physiological characteristics, have thus far been regarded as constants. As the observed variation in internal doses could not be explained by variation in intake and constitution alone, it had to be assumed that the toxicokinetic parameters contain significant interindividual variability as well. This conclusion was supported by the following observation. The Monte Carlo analysis evaluating variation in intakes and organ weights resulted in variation in concentrations in both kidney and liver. However, between individual runs the ratio of these concentrations appeared to be constant, due to the fact that the model practically behaves linearly at the actual exposure level of the population. The observed kidney-to-liver ratios, on the other hand, did show considerable variation between individuals (CV = 58%) (7).

To incorporate physiological variation into the toxicokinetic model, we assumed lognormal distributions for at least some of the model's parameters. Thus, each of these stochastic parameters can be represented by two substitute parameters: a mean and a variance (on log-scale). The question now is: How can these substitute parameters, in particular, the variances, be assessed? In the deterministic version of the model, the unknown parameter values can be obtained by calibration using standard methodology (optimization of criterion function, e.g., sum of squares or likelihood). However, to date no formal methodology is available to estimate variances of parameters in dynamic models by calibration to data. Therefore we adopted the approach of trial and error, providing the parameters with variances such that the observed variation in the data is satisfactorily predicted by the model. Unfortunately, there is no unique solution; the effect of increasing the variance in one parameter can easily be cancelled by decreasing the variances of others. Therefore, a limited number of stochastic parameters was assumed. In selecting these parameters we proceeded as follows.

First, we subjected the model to a sensitivity analysis with respect to the kidney-to-liver ratio. This was done by giving each parameter a uniform distribution with a range of 5% of the nominal value. Then, 1000 randomly drawn parameter combinations were evaluated. Multiple regression of the model output on the parameters resulted in three conspicuous regression coefficients, indicating the associated parameters to be clearly more sensitive than the others, viz., fraction of absorbed cadmium bound to metallothionein in the blood, transport rate from liver to metallothionein in blood, and urinary excretion rate [denoted as C7, C14, and C19, respectively, in Nordberg and Kjellström (6)]. These parameters were all provided with CVs of 45%. At these values the Monte Carlo

Figure 2. Illustration of the Monte Carlo analysis. Panel (A) shows a random sample of functions representing cadmium absorbed from food (and smoking, if applicable), (B) denotes the toxicokinetic model with a number of parameter distributions, (C) gives a sample of organ weight functions, and (D) shows the resulting output functions; in this case, the cadmium concentration in the kidney. In each Monte Carlo run a single individual, exemplified by the thick curves and dots, is evaluated. Note that each individual is simulated over a lifetime.
analysis predicted a variation in kidney-to-liver ratios similar to that observed. However, in this situation the predicted concentrations in both kidney and liver were still less variable than the observed concentrations. Therefore, we once more subjected the model to a sensitivity analysis, but now with respect to the concentrations in kidney and liver. This appointed another parameter (absorption rate from food) as sensitive. Providing this parameter with a CV of 65% (and decreasing the CVs of the other three parameters to 40%) resulted in agreement between predicted and observed variation in concentrations in kidney and liver as well as in the kidney-to-liver ratio. Figure 2 illustrates the complete Monte Carlo procedure.

We have shown how the presumed sources of variation were traced and quantified, with the result that the predicted variation in internal concentrations agreed with the observed variation. This effort enables us to assess the risk for a particular group. Some examples of risk factors are smoking, anemia (8), and living in a polluted area (7). For instance, anemia increases the absorption of cadmium from food. We incorporated this into the model by assuming that the absorption of cadmium in anemic women (at age 14-45 years) is increased by a factor of 2 (median), with a CV of 25% to represent interindividual variability (8). Also, combinations of these risk factors may be evaluated. For any given risk group the Monte Carlo analysis can be completed by constructing histograms of the maximum cadmium concentration in the kidney achieved during a lifetime in each individual run. Figure 3 shows the histograms thus resulting for several risk groups. Comparing these histograms with the critical value (200 mg/kg fresh weight in the kidney cortex) at which effects are expected (5) results in the fraction of the population at risk. From these distributions it can be concluded that at present the risk of exceeding the critical value is very small in the Dutch population. In the case of a special subgroup, where all known risk factors were combined (Fig. 3c), the probability of exceeding the critical value was estimated at several per thousand.

The present case study illustrates a methodological framework for evaluating interindividual variability in intake and toxicokinetics, resulting in predicted distributions of internal doses. In the case of cadmium where the critical internal dose is relatively well known, these distributions immediately result in an estimate of the fraction of the population at risk or the probability of adverse effects for a particular combination of risk factors.

Basically, the approach followed here is not unusual in risk assessments. It amounts to calibrating a model to a set of data and subsequently using the model for extrapolation purposes. The difference is that the model not only intends to describe the average individual but the population as a whole (i.e., including nonaverage individuals). One might speak of second-order as opposed to first-order modeling. In first-order modeling the distribution of internal dose as a function of age may be obtained by adding the estimated residual variation to the output function. A danger of this approach is that when data are not equally available at all ages, the residual variation may wrongly be taken to be homogeneous. Therefore, percentiles (risks) at ages with no or poor data may be misestimated. In second-order modeling, the relation of the variability in internal doses with age follows from the model itself. The important advantage of second-order modeling is that it allows for the evaluation of specific risk factors such as smoking or anemia, assuming that the variation in other factors (e.g., in intake), remains unchanged. Moreover, risk factors themselves can be evaluated in a statistical fashion. For example, instead of evaluating the "typical" anemic female, we used distributions representing variation in the severity of the anemia as well as variation in duration and age at which it may occur. Smoking habits in the population can also be evaluated statistically. Such evaluations are not possible in a first-order model. Furthermore, the estimation

---

**Figure 3.** The distribution of predicted maximum cadmium concentrations in the kidney over a lifetime for a number of risk groups. (A) Nonsmoking males, (B) smoking females, (C) worst-case scenario: smoking, anemic females living in Kempenland (a cadmium-contaminated area in the Netherlands). Arrows: the critical value at which effects are expected, assuming a ratio of 1.5 between kidney cortex and whole kidney concentration (10).
of distributions of internal doses from a first-order model will fail when the model is nonlinear, or nonlinear in a certain input/parameter range. The latter is the case in the present cadmium model. Indeed, the subgroup histograms of Figure 3 appear to have different CVs.

Although we believe that the general framework presented here is promising, the trial-and-error approach of assessing the parameters’ variances is obviously unsatisfactory. It is unclear whether some particular allocation of variances to the parameters other than the allocation we used might result in identical predictions of internal dose distributions. A practical objection is the rather time-consuming procedure. These considerations call for a formal procedure for calibrating stochastic parameters in dynamic models. To date, a maximum likelihood procedure has only been worked out for a simple elimination model with a single stochastic parameter (9), which proved to be rather tedious. For more complicated models, numerical procedures are necessary. At present we are trying to develop such procedures.

One may suspect that calibrated parameter variances will tend to be correlated, especially since data on internal doses are usually not equally available at all ages. Fortunately, a promising development has taken place with respect to toxicokinetic modeling: the introduction of physiologically based pharmacokinetic (PBPK) models. In this class of models the parameters have physiological correlates. This offers, at least in principle, the opportunity to measure the parameters of PBPK models, including their interindividual variation. At present we are investigating the possibilities of quantifying the variation in parameters of a PBPK model in a case study on dioxins using published measurements.

An interesting detail in the present analysis concerns the kidney-to-liver ratio. Being constant with respect to variation in intake and constitution, this ratio offered an opportunity to discriminate between these sources of variation on the one hand and variation in model’s parameters on the other. This property of nearly linear models can be exploited in other cases as well.

The primary advantage of the approach presented in this paper is the possibility for estimating risks in particular groups. A second advantage is the ability to evaluate future risks and the effects of policy measures. For example, the small risks that were assessed in the present analysis even for extreme risk groups may increase in the future if no measures are taken. When an environmental model is available relating the emission of cadmium to concentrations in soil and foods (and other relevant routes), changes in emissions can be evaluated in terms of risks. These uses make it worthwhile to technically improve the method so that it can indeed be used for risk management purposes.

References
1. Travis CC, Arms AD. The food chain as a source of toxins exposure. In: Toxic chemicals, health and environment (Lave LB, Upton AC, eds). New York:Prager, 1987:95–113.
2. Slob W. Modeling long term exposure of the whole population to chemicals in food. Risk Anal 13:525–530(1993).
3. Dutch Ministries of WVC and LNV, VCP: Wat eet Nederland; resultaten van de voedselconsumptiepeiling 1987–1988 [in Dutch]. Rijswijk, the Netherlands: Dutch Ministries of WVC and LNV, 1988.
4. Huishof KFAM, Van Staveren WA. The Dutch national Food Consumption Survey: design, methods and first results. Food Policy 16:257–260(1991).
5. WHO. Thirty-third report of the Joint FAO/WHO Expert Committee on Food Additives. Evaluation of certain food additives and contaminants. WHO Technical Report Series 776. Geneva:World Health Organization, 1989.
6. Nordberg GF, Kjellström T. Metabolic model for cadmium in man. Environ Health Perspect 28:211–217(1979).
7. Kreis IA. Health effects of cadmium contamination in Kempenland (PhD thesis): Leiden, the Netherlands:Rijksuniversiteit, 1992.
8. Flanagan PR, McLellan JS, Haist J. Increased dietary cadmium absorption in mice and human subjects with iron deficiency. Gastroenterology 74:841–846(1978).
9. Bedaux JJM, Kooijman SALM. Stochasticity in deterministic models. In: Handbook of statistics, vol 12. Environmental statistics (Patil GP, Rao CR, eds). Amsterdam:Elsevier Science Publishers (in press).
10. Livingston HD. Measurement and distribution of zinc, cadmium, and mercury in human kidney tissue. Clin Chem 18:67–72(1972).