Uncertainty and Variability Modelling of Chemical Exposure through Food

by

Verdonck FAM^{1,5}, Sioen I^{2,3}, Baert K³, Van Thuyne N^{1,2,3}, Bilau M², Matthys C², De Henauw S², De Meulenaer B³, Devlieghere F³, Van Camp J³, Vanrolleghem PA¹, Van Sprang P⁵, Verbeke W⁴, Willems J²

Abstract

The deterministic approach in current EU risk assessment directives of new and existing substances deals with uncertainty as a technical construct under the format of worst-case assumptions and safety factors as if these 'certain uncertainties' can lead to certain risk estimates. In this way, the considered uncertainty and variability in the exposure and risk estimates is made insufficiently transparent which prevents policymakers to properly assess and manage potential chemical risks. Uncertainty and probabilistic analysis is a useful process for risk assessors and managers that quantifies to the extent possible the uncertainty and variability of chemical exposure to humans through food.

Keywords: Monte Carlo method, risk assessment, uncertainty

¹ Ghent University, Department of Applied Mathematics, Biometrics & Process, Ghent, Belgium

² Ghent University, Department of Public Health, Ghent, Belgium

³ Ghent University, Department of Food Safety and Food Quality, Ghent, Belgium

⁴ Ghent University, Department of Agricultural Economics, Ghent, Belgium

⁵ EURAS, Ghent, Belgium

Correspondence address: Frederik Verdonck, Coupure links 653, 9000 Gent, Belgium. E-mail: frederik.verdonck@biomath.ugent.be

Introduction

Environmental pollution by toxic substances has led to regulations on the production and use of chemicals. The tools used for assessing the potential impact of chemicals to human health are risk-based. Usually, such a risk analysis process can be subdivided into three interconnected components (terminology according to the EC (1):risk assessment (hazard identification, hazard characterisation or effects assessment, exposure assessment, risk characterisation), risk management (process used for deciding between policy alternatives in consultation with stakeholders e.g. mitigation, cost-benefit etc), and risk communication (interactive exchange of information and opinions concerning risk and risk management with all involved stakeholders). The basic approach for risk assessment is to compare an exposure estimate with a threshold for effect (e.g. Tolerable Daily Intake or TDI) covering certain adverse human health effects.

The European risk assessment principles for new and existing chemicals are laid down in respectively Commission Directive 93/67/EEC (2) and Commission Regulation (EEC) 1488/94 (3) and are described in detail in the Technical Guidance Documents (4) and the software package EUSES (5). Recently, the European Commission adopted a legislative proposal for implementing a new EU chemicals policy, known as REACH (Registration, Evaluation and Authorisation of Chemicals) (6). Populations considered in the current and upcoming legislation are on the one hand consumers and workers, both exposed directly during or following use of chemicals, and on the other hand humans exposed to chemicals following transport to and distribution in the environment (indirect exposure). Exposure via the environment is the focus of this paper. The European regulation (EC) 178/2002 states that food law is also based on risk analysis, except where this is not appropriate to the circumstances or the nature of the measure (7).

The current exposure assessment schemes are largely deterministic and uncertainty and/or variability issues are accounted for by means of cautionary measures which are implicitly embedded in calculation schemes and rules (e.g. use of safety factors, choice of 90th percentile values, worst-case assumptions) (8). Consequently, this makes the exposure assessment insufficiently transparent to a risk manager or decision-maker as to how conservative or realistic the exposure and risk estimate actually are. There is, therefore, a need for a proper and transparent treatment of sources of uncertainty and variability during the exposure and risk assessment process. The goal of this paper is to demonstrate that probabilistic techniques allow for such increased transparency of variability and uncertainty in the chemical exposure assessment through food. Increased transparency improves the subsequent risk assessment and management process. An example on the dioxin intake through seafood consumption by the Belgian population will be used as an example throughout the paper.

Current exposure assessment

Indirect exposure of humans via the environment may occur by consumption of food (fish, crops, meat, milk...) and drinking water, inhalation of air and ingestion of soil. The different routes of exposure are illustrated in Figure 1. Exposure via soil ingestion and dermal contact is not addressed in this EU guidance because soil ingestion and dermal contact represent significant exposure routes only for specific situations of soil pollution. The indirect exposure is assessed by estimating the total daily intake of a substance based on the predicted or measured environmental concentrations for (surface) water, groundwater, soil and air (4).





The assessment of the indirect exposure via the environment is carried out following the sequence of the step-wise procedure (4):

- assessment of the concentrations in intake media (food, water, air and soil);
- assessment of the intake rate of each medium;
- combination of the concentrations in the media with the intake of each medium (and if necessary/available, using a factor for the bioavailability through the route of intake).

The calculation methods described are simple, flexible and "state of the art" methods for predicting the indirect exposure to humans via the environment. The concentrations in the environmental compartments which are required as input data in the models for the calculation of the total daily intake via the different exposure routes should be derived on the basis of monitoring data and/or modelling. The concentration of a substance in food is related to its concentration in water, soil and air and is also dependent on its bioaccumulation or biotransfer behaviour. The models for the estimation of daily intake allow using local or regional environmental concentrations, as appropriate (4).

Sources of uncertainty and/or variability are accounted for in a worstcase approach (e.g. 90th percentile or maximum of concentration data, worst-case consumption level, use of safety factors). This typically leads to an overestimated, worst-case risk estimate. For example, a worstcase exposure to dioxin of 8.9 pg TEQ/d/kg BW is obtained by multiplying a worst-case dioxin concentration in salmon of the Baltic sea of 7.3 pg TEQ/g ww with a worst-case salmon consumption of 600 g/week and an average Body Weight (BW) of 70 kg. In reality, human behaviour shows an appreciable amount of variation between the different EU countries. But also within countries, large variations occur between individuals. As a consequence, indirect exposure will vary greatly over the population to be protected. A more transparent approach is to explicitly consider this inter-individual variability and related uncertainty.

Probabilistic assessment

Single point or deterministic modelling involves using a single "best guess" or "worst case" estimate of each parameter within a model to determine the model's outcome(s) (see above example). "What if" scenarios can then be conducted to determine how much the model's outcome changes when different parameters or assumptions are considered (e.g. increasing or decreasing the consumption). Probabilistic exposure assessment is similar to "what if" scenarios in that it generates a number of possible scenarios. However, it goes one step further by effectively accounting for every possible value that each parameter could take and weighting each possible scenario by the probability of its occurrence. In this way, insight is provided into the uncertainty and variability associated with exposures. Probabilistic exposure assessment achieves this by defining each parameter within a model by a probability distribution.

First, the definitions of uncertainty and variability are given. It is subsequently illustrated that in the scientific community, there is an increasing awareness on the usefulness of more explicit consideration of uncertainty and variability. Then, an overview is given on the characterisation and propagation of uncertainty and variability.

Uncertainty and variability

The distinction between uncertainty and variability is considered important to make (11, 9, 10). Note that in literature, uncertainty is sometimes also used as an umbrella term for both variability and limited or no knowledge. Variability represents inherent heterogeneity or diversity in a well-characterised population. Fundamentally a property of nature, variability is not reducible through further measurement or study. In the exposure assessment of human health risk, the variability in lifestyle, diet and consumption pattern of the human population is an important source next to the spatial (and temporal) variability of the contamination of food items and the differences in duration and in route of exposure. Uncertainty represents limited knowledge, comprising (ir)reducible ignorance, inexactness (e.g. measurement errors), lack of observations/measurements (sampling uncertainty), indeterminacy or lack of perfect information about phenomena or models (e.g. model structure uncertainty), and can partly be reduced through further research. Sources of uncertainty in exposure assessment are situated at the level of the concentration data: sampling uncertainty (representativeness of time period and place of sampling), uncertainty on the origin of the food item; at the level of the intake rate: uncertainty on the accuracy of the intake (problem of over- and underreporting), consumer's claimed versus actual behaviour, uncertainty of the representativeness of the study sample for the overall population; and at the level of effects/risk characterization: e.g. from animal to human toxicity extrapolation. In practice, only the quantifiable or known sources of uncertainty are typically considered for pragmatic reasons. It is therefore important that the risk assessor should be aware of the presence of ignorance and unquantifiable sources of uncertainty.

A parameter can be characterised by both uncertainty and variability. In this case, the parameter is called a second-order parameter and is represented by probability distributions in two dimensions. For each percentile of the variability distribution, an uncertainty or confidence interval can be calculated (i.e. the uncertainty distribution). In Figure 2, the variability of a parameter is presented as a cumulative distribution function. The uncertainty distribution can also be represented by a cumulative or density distribution function; however, for communication purposes, it is often represented by a 90% uncertainty or confidence interval or band (grey bands in Figure 2).





State-of-the-art

The history of the use of probabilistic techniques is long and varied. In a major advance during the 1940s, the Monte Carlo simulation, a random sampling technique for solving difficult deterministic equations, originated from the work of Ulam, von Neumann, and Fermi (11). In recent years, there has been growing interest in the application of probabilistic techniques to the estimation of exposure to food chemicals. In 2000, a joint FAO/WHO Workshop was held on Methodology for Exposure Assessment of Contaminants and Toxins in which probabilistic techniques were proposed for exposure assessment (12). Recently, the Interdepartmental Group on Health Risks from Chemicals provided guidance on good exposure assessment practice for human health effects. This guidance comprises how to model exposure concentration, how to take into account variability, as well as how to conduct an uncertainty and sensitivity analysis and how to evaluate the results of an exposure assessment (13). Prompted by the growing interest in the area of probabilistic modelling for exposure assessments, the Monte Carlo project on the 'Development, validation and application of stochastic modelling of human exposure to food chemicals and nutrients' was funded under the EC Fifth Framework Programme in 1999-2002 (14, 15). The output of the project includes custom-built software for modelling food chemical exposure, reports on data input issues, data on pesticides residues and nutrients, and the production of practical guidelines for probabilistic modelling of food intakes.

A lot of scientific research papers were published describing the application of probabilistic modelling in the field of food safety in order to evaluate the intake of contaminants via food ingestion. Some examples are presented here. Vermeire et al. (5) shows the advantages and possibilities of a probabilistic human health risk assessment of the industrial chemical dibutylpthalate (DBP), including the uncertainty in the exposure estimate, in the effect parameters and in the assessment factors used in the extrapolation from experimental animals to sensitive human beings. Vrijens et al. (15) demonstrates a probabilistic intake assessment of PCDDs. PCDFs and dioxin-like PCBs in the so-called 1999 'Belgian dioxin incident'. For this, Monte Carlo simulation techniques were used to combine detailed 7-day food intake data on the individual level with 'background' and 'incident-related' food contaminant data. Tressou et al. (17) presents the use of extreme value theory in probabilistic exposure assessment to heavy metals (lead, cadmium, mercury) via seafood consumption. Joao Paulo et al. (18) presents a Bayesian method which incorporates multivariate modelling of food consumption and modelling of pesticide measurements which are for a large part below a measurement threshold. It was shown that Bayesian modelling compares well with an empirical Monte Carlo modelling in a data-rich situation. All these examples highlight beneficial aspects of a probabilistic exposure assessment.

Uncertainty and variability characterisation

The outcome of an exposure/intake model is often mostly influenced by specific input parameters and assumptions (e.g. how does the selection of a probability distribution influence the exposure assessment?). To identify the critical parameters or assumptions (e.g. specific exposure pathways or certain assumptions with respect to input parameters) that drive the exposure model and consequently to efficiently allocate resources for further data collection, sensitivity analysis can be used. The basic approach is to allow for a subset of the input parameters to vary within prescribed ranges and to determine how much the model output (e.g. exposure) changes in response to changes in the values for each input parameter. Of the several variants of this basic approach for sensitivity analysis that are available, no single approach will serve as the best analysis for all modelling efforts. For example, sensitivity ratios can be used where the ratio is equal to the percentage change in output divided by the percentage change in input for a specific input parameter. Risk or intake estimates are considered most sensitive to input parameters that yield the highest ratios. Guidance on how to perform a sensitivity analysis is provided by Saltelli et al. (19) or Cullen & Frey (11). An example is given in Figure 3 on the dioxin intake through seafood consumption. Parameters are arrayed from most influential at the top to least influential at the bottom. It can be observed that mainly the variation in dioxin concentration of the salmon and the herring contributes to the variation in the resulting total dioxin intake. The consumption levels are less influential.

If an input parameter is found to be important to the outcome, additional data can be collected (if not already available) to characterise its variability and/or uncertainty. A number of steps then need to be conducted in order to characterise the uncertainty and variability of an input parameter. First, a number of data points or, if lacking, expert knowledge is needed to quantify its uncertainty and variability. Second, it can be checked whether a parametric or nonparametric distribution is more appropriate for the input parameters. Parametric methods assume that the data come from a fixed underlying distribution. This assumption enables to work with smaller sample sizes. Nonparametric methods rely on the data points themselves. This makes them less vulnerable to deviations from certain distribution assumptions but more vulnerable to deviations in the data points. Graphical plots like histograms and scatterplots can be used to explore the data of the parameter and help in selecting a parametric or nonparametric distribution. A parametric distribution could be selected based on best fitting criteria (such as goodness-of-fit tests) and expert knowledge. The parametric or nonparametric distribution can then be characterised. For example, a lognormal (parametric) distribution was selected to represent the spatial variability of the dioxin concentration (20) in salmon from the Baltic sea (see lognormal fit in Figure 4). A third step involves exploring whether the parameter under investigation is considered to be uncertain, variable or both. Figure 3. Example outcome of a sensitivity analysis on the dioxin intake (pg TEQ/ kg bw/d) through seafood consumption (conc: concentration, consump: consumption)



Sensitivity to Dioxin.intake

An example of the latter is the dioxin concentration in salmon that varies between locations (spatial variability) but is also uncertain due to a limited number of samples. If a parameter is found to be uncertain and variable at the same time, its uncertainty and variability can be characterised as illustrated in Figure 4 for dioxin concentration in salmon. In this example, the bootstrapping technique was used to separate variability from uncertainty (under the assumption of a lognormal distribution). Given a sample of size *n*, the general approach in bootstrap simulation is to use the data points themselves or to assume an underlying distribution, to perform *r* replications (e.g. r = 5,000) of the original data set by randomly drawing, with replacement, *n* values from the data points or the underlying distribution, and then calculate *r* values of the statistic of interest. These can be used to determine the uncertainty distribution of the statistic of interest (21). More information on selecting and fitting

distributions and the characterisation of variability and sampling uncertainty can be found in literature (11, 22, 23, 24).

Once all parameters in the exposure and effects assessment are characterised as either point estimates or probability distributions, it should be checked whether there are correlations between these parameters. This is important for subsequent uncertainty and variability propagation, where the different parameters are combined. Vose's (24) 'cardinal rule of risk analysis modelling' is 'every iteration of a risk analysis model must be a scenario that could physically occur'. For example, the high body weight can be highly correlated with the consumption level of fatty food products. Further guidance can be found in Vose (24).

Uncertainty and variability propagation

Once the most important input parameters are identified and their uncertainty and variability is quantified, the uncertainty and variability of the input parameters can be propagated through the intake model in order to obtain an exposure estimate. There are a variety of ways to propagate information about uncertainty or variability through the model (11). Here, the Monte Carlo analysis is selected. In Monte Carlo analysis, random samples of model input parameters are selected according





to their respective assigned probability distributions. Once the samples from each input distribution are selected, the set of samples is entered into the model. The model is then solved as one would do for any deterministic analysis. Each individual model result is stored and the process is repeated until the specified number of model iterations is completed. Instead of obtaining a discrete number of model results (as in a deterministic simulation), a set of output results is obtained, which are all together used to characterise an output distribution (11).

A first order or one-dimensional Monte Carlo simulation can only propagate variability or uncertainty, but not both at the same time without having difficulties to interpret the output. It is therefore recommended that for this, among other propagation techniques, a second order or two-dimensional or embedded Monte Carlo simulation be applied (11). It consists simply in two Monte Carlo loops nested one inside the other. The inner loop deals with the variability of the input variables, while the outer one deals with uncertainty. For each shot of a (uncertain) parameter value in the outer loop a whole distribution is created in the inner loop based only on variability (see Figure 5 and Figure 6). In this way changes in variability-dependent frequency distributions under the influence of parameter uncertainty can be quantified.



Figure 5. Simulation algorithm of a second order Monte Carlo simulation





The final result after characterisation and propagation of uncertainty and variability through a food intake model is an intake probability distribution with an uncertainty or confidence band as illustrated in Figure 2. The intake probability distribution transparently communicates how variable the intake and risk of a chemical substance is among the human population under study. The uncertainty band transparently communicates the level of uncertainty to the risk managers. It quantifies how reliable the exposure assessment was. This improves risk management. Furthermore, this approach will allow communicating more effectively and more efficiently about risk, including risk uncertainty and variability, which will be a key to building consumer trust (25).

Not all sources of uncertainty and variability can be quantified. Model uncertainty and decision rule uncertainty for example are difficult to quantify and propagate through the assessment. These remaining sources of uncertainty and variability could in this case be added to the assessment on a (semi-)qualitative basis to the extent possible.

Conclusions

In current EU risk assessment directives of new and existing substances, uncertainty is insufficiently made transparent in the exposure and risk estimates preventing policy-makers to properly assess, manage and communicate about potential chemical risks. Uncertainty and probabilistic analysis is a useful process for risk assessors and managers that quantifies to the extent possible the sources of uncertainty and variability of human exposure.

Acknowledgement

This research was funded by a scholarship from the Flemish Institute for the Improvement of Scientific-Technological Research in the Industry (IWT), and the Belgian Science Policy through the SPSD II project CP/02/56. The authors would also like to thank the two anynomous referees for their useful suggestions to improve the paper.

Samenvatting

De deterministische aanpak in de huidige EU-risicoanalyserichtlijnen van bestaande en nieuwe stoffen behandelt onzekerheid als een technisch construct onder de vorm van worst-case veronderstellingen en veiligheidsfactoren alsof deze 'zekere onzekerheden' kunnen leiden tot zekere risicoschattingen. Op deze manier worden de onzekerheid en variabiliteit in blootstelling- en risicoschatting onvoldoende transparant gemaakt. Dit verhindert beleidsmakers om behoorlijk potentiële risico's te analyseren en te beheren. Onzekerheids- en probabilistische analyse is een nuttig proces voor risicoanalisten en -managers dat, in de mate van het mogelijke, de onzekerheid en variabiliteit van chemische blootstelling aan mensen door voeding kwantificeert.

Résumé

L'approche déterministique dans les directives UE actuelles d'évaluation des risques pour substances nouvelles et existantes traite l'incertitude comme une construction technique qui est basée sur l'utilisation de suppositions pire cas et de facteurs de sûreté comme si ces 'incertitudes certaines' peuvent mener à des estimations certaines des risques. De cette façon, l'incertitude et la variabilité considérées dans les estimations de l'exposition et du risque sont insuffisamment transparent ce qui empêche le décideur d'évaluer et de gérer convenablement les risques potentiels. L'analyse d'incertitude et probabilistique est un procédé utile pour les contrôleurs et gestionnaires de risques et aides à quantifier l'étendue de l'incertitude et de la variabilité dans l'exposition de l' homme suit à l'ingestion de nourritures contaminées.

References

- 1. EC (European Commission). Communication from the commission on the precautionary principle. COM 2000; 1.
- EC (European Commission). Commission Directive 93/67 EEC of 20 July 1993, laying down the principles for the assessment of risks to man and the environment of substances notified in accordance with Council Directive 67/548/EEC. Official Journal of the European Communities L227. 1993.
- EC (European Commission). Commission Regulation (EC) 1488/94 of 28 June 1994, laying down the principles for the assessment of risks to man and the environment of substances notified in accordance with Council Directive 67/548/EEC. Official Journal of the European Communities L161. 1994.

- EC (European Commission). Technical Guidance Document in support of Commission Directive 93/67/EEC on risk assessment for new notified substances and Commission Regulation No. 1488/94 on risk assessment for existing substances. Office for Publications of the European Communities. 1996.
- Vermeire T, Jager T, Janssen G, Bos P, Pieters M. A probabilistic human health risk assessment for environmental exposure to dibutylphthalate. Hum Ecol Risk Ass 2001; 7(6): 1663-79.
- CEC. Proposal for a regulation of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restrictions of Chemicals (REACH) establishing a European Chemical Agency and amending Directive 1999/45/EC and Regulation (EC) 'on Persistent Organic Pollutants'. In COM 2003; 644.
- EC (European Commission). Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety. Official Journal of the European Communities L31. 2002.
- 8. Verdonck FAM, Van Sprang P, Vanrolleghem PA. Uncertainty and precaution in European environmental risk assessment. Wat Sci Tech 2005; 52(6): 227-34.
- Hoffman F and Hammonds F. Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Anal 1994; 14(5): 707-12.
- 10. van Asselt MBA, Rotmans J. Uncertainty in integrated assessment modelling: from positivism to pluralism. Climate Change 2002; 54: 75-105.
- Cullen A and Frey H. Probabilistic techniques in exposure assessment. A handbook for dealing with variability and uncertainty in models and inputs. Plenum, New York 1999: 335.
- 12. WHO. Methodology for exposure assessment of contaminants and toxins in food. www.who.int/fsf. 2000.
- IGHR (Interdepartmental Group of Health Risk for Chemicals). Guidelines for good exposure assessment practice for human health effects of chemicals. http://www.le.ac.uk/ieh/pdf/CR10.pdf. 2004.
- 14. Gibney MJ and van der Voet H. Introduction to the Monte Carlo project and the approach to the validation of probabilistic models of dietary exposure to selected food chemicals. Food Add Cont 2003; 20Suppl(1): S1-7.
- McNamara C, Naddy B, Rohan D and Sexton J. Design, development and validation of software for modeling dietary exposure to food chemicals and nutrients. Food Add Cont 2003; 20Suppl(1): S8-S26.
- Vrijens B, De Henauw S, Dewettinck K, Talloen W, Goeyens L, Willems JL. Probabilistic intake assessment and body burden estimation of dioxin-like substances in background conditions and during a short food contamination episode. Food Add Cont 2002; 19(7): 687-700.
- Tressou J, Crépet A, Bertail P, Feinberg MH, Leblanc JCh. Probabilistic exposure assessment to food chemicals based on extreme value theory. Application to heavy metals from fish and sea products. Food Chem Tox 2004; 42: 1349-58.
- Joao Paulo M, van der Voet H, Janssen MJW, ter Braak CJF, van Klaveren JD. Risk assessment of dietary exposure to pesticides using a Bayesian method. Pest Manag Sci 2006 [in press].
- Saltelli A, Chan K, Scott EM. Sensitivity Analysis. Wiley Series in Probability and Statistics. ISBN 0-471-99892-3, 2000: 486.

- 20. Sioen I, Van Camp J, Verdonck F, Van Thuyne N, Willems J, De Henauw S. How to use secondary data on seafood contamination for probabilistic exposure assessment purposes? Main problems and potential solutions. HERA [in press].
- Verdonck FAM, Jaworska J, Thas O & Vanrolleghem PA. Determining environmental standards using bootstrapping, Bayesian and maximum likelihood techniques: a comparative study. Anal Chim Acta 2001; 446 (1-2): 429-38.
- EPA. Report of the workshop on selecting input distributions for probabilistic assessments.
 U.S. Environmental Protection Agency, Washington, DC, 1999. EPA/630/R-98/004.
- EPA. Risk assessment guidance for superfund (RAGS): Volume III Part A. Process for conducting probabilistic risk assessment. U.S. Environmental Protection Agency, Washington, DC. 2001.
- 24. Vose D. Quantitative risk analysis. A guide to Monte Carlo simulation modeling. John Wiley & sons. 1996. 317.
- Frewer L, Fischer A, Scholderer J, Verbeke W. Food safety and consumer behaviour. In: Jongen W., Meulenberg M (Eds.), Innovation in agri-food systems: Product quality and consumer acceptance, Wageningen: Wageningen Academic Publishers, ISBN 90 769 986 55. 2005: 125-46.