

This is an electronic preprint version of the article

HEINZL,H., MITTLBÖCK,M. & EDLER,L. (2005). On the role of ex post uncertainty assessment for risk management. *International Journal of Risk Assessment and Management*, **5**, 206 - 215.

published in

International Journal of Risk Assessment and Management © 2005
Inderscience Enterprises Ltd.

<http://www.inderscience.com/ijram>

On the role of *ex post* uncertainty assessment for risk management

Harald HEINZL¹, Martina MITTLBOECK² and Lutz EDLER³

¹) Core Unit for Medical Statistics and Informatics, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria; e-mail: harald.heinzl@meduniwien.ac.at; fax number: +43 (0) 1/40400-6687

²) Core Unit for Medical Statistics and Informatics, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria; e-mail: martina.mittlboeck@meduniwien.ac.at; fax number: +43 (0) 1/40400-2278

³) Division of Biostatistics, German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany; e-mail: edler@dkfz.de; fax number: +49 (0) 6221/42-2397

Abstract: A risk management decision whether a chemical compound present in the environment is potentially hazardous to human health will be among others based on large-scaled statistical analyses of empirical data of various sources. Observations from highly exposed and well-documented occupational cohorts are usually a matter of particular interest. Due to scientific progress, various new aspects will inevitably emerge in the course of time. These new aspects are not always in full accordance with the assumptions of the original statistical analyses so that concern may evolve whether the original foundation of the risk assessment is still valid in the main. We suggest to account for such doubts by an *ex post* uncertainty assessment of the original statistical analyses. The case of the risk assessment whether 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) is a human carcinogen is used as example. We call for the scheduling of future *ex post* uncertainty assessments already at the time of the risk management decision.

Key words: Risk assessment, uncertainty assessment, dioxin, environmentally hazardous compound, cancer, computer simulation study, Carrier kinetic.

1. Introduction

Concerns on risks either caused through natural or through man made substances, released unintentionally or accidentally into the environment, has increased the number as well as the extent and thoroughness of risk assessments over the last decades. In this course, one has realized that the process of risk assessment bears more on uncertainty than what can be covered under the notion of statistical variability and that uncertainty can arise at all levels of this process. Uncertainty is prevalent in the segment of hazard identification as well as in the segment of exposure assessment thereby impairing dose estimation. Even more, its effects may be increased further by intrinsic uncertainty in the process of the dose-response modelling itself, e.g., caused by limited knowledge about the functional dose-response relationship. At the end, this uncertainty is propagated to hazard characterisation and risk quantification and as such it has the potential to infiltrate the risk management decision. Consequently, assessment of uncertainty must be considered as an integral part of risk assessment.

However, the case of uncertainty is usually not closed with a specific risk management decision. There may be cases where the risk management decision will be almost immediately questioned during a review process and where a formal, independent and *immediate* uncertainty assessment is required by some stakeholders. More important and more challenging from a scientific perspective, however, is the situation when newly emerging scientific aspects may seriously question the foundations of the original risk assessment at a later date. Obviously, there will be a need to address such novel aspects by an uncertainty assessment as well. In order to point to the inevitably inherent timely delay we want to denote

such an assessment *ex post uncertainty assessment*, see Figure 1. The goal of the *ex post* uncertainty assessment is to provide an evaluation of the limits of the available statistical analyses under the light of the newly emerging scientific aspects, or in other words, the uncertainty assessment should "*tell us how much we can be wrong and still be fine*" (Bois and Diack 2005).

The carcinogenic risk assessment of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD, "Seveso-dioxin") is used for the sake of illustration in the following as it provides a well known example of an important risk management problem. Highly exposed occupational cohorts produced important evidence for the evaluation of the International Agency for Research on Cancer (IARC) that TCDD is carcinogenic to humans (IARC 1997, McGregor et al. 1998). One of these cohorts is the so-called Boehringer cohort whose data have undergone careful and elaborate large-scaled statistical analyses, see Becher et al. (1998) and references therein. Since then several novel scientific aspects have been raised which directly relate to the methods and underlying assumptions of the original statistical analyses. This motivated an *ex post* uncertainty assessment of the above intention. Section 2 exemplarily presents a small but representative part of this assessment and discusses its implications. Section 3 generically establishes the role of *ex post* uncertainty assessment.

2. Example: *Ex post* uncertainty assessment for the Boehringer cohort study

2.1 Human data

The Boehringer cohort comprised more than 1500 workers mainly engaged in the production of herbicides from 1950 until 1984 in Germany (Flesch-Janys et al. 1995, Becher et al. 1998). Some twenty different working areas were identified in that plant and clustered into five main working areas of different TCDD exposure levels. There was e.g. one working area only

existing in the 1950ies which seemed to have been extremely contaminated. Those workers were at least 20-times as much exposed than all other workers (Becher et al. 1998). When TCDD concentration measurement in humans became accomplishable, the levels of TCDD in blood serum or body fat samples were determined in 245 workers between 1985 and 1994.

2.2 Methods

Cancer is considered as a slowly evolving disease. Therefore, an increase in cancer incidence due to an increase of the TCDD body burden of those workers may manifest itself not after some latency period. This fact needs careful accountance when a dose-response relationship is to be established. Toxicokinetic models are capable to reconstruct TCDD exposure concentrations during a worker's lifetime such that time-dependent TCDD dose metrics like the area under the concentration-time curve (AUC) can be constructed on an individual basis. TCDD concentration-time curves can be obtained by combining job exposure matrices and individual workplace history data. Exposure estimates can be obtained using the back-calculation method, that is, available TCDD measurements are related to the corresponding workplace history data in a statistical regression model (Becher et al. 1998). Potentially rather high TCDD exposures, particularly occurring in the 1950ies, can so be reconstructed from generally lower exposures determined in the 1980 and 1990ies. This reconstruction is a rather delicate task from a statistical viewpoint as an extrapolation beyond the observed range of measurements takes place. Notice, extrapolated results cannot be verified with common model checking techniques (e.g. residual analysis). The validity of the results rises and falls with that of the underlying model assumptions.

A one-compartmental mass-balance equation for describing the amount $A(t)$ of TCDD in the body at time t ,

$$dA(t)/dt = intake(t) - elimination(t),$$

was adapted for use with the Boehringer cohort data. Since available data did not support more complex approaches, it was assumed that the total lipid volume (TLV) of the body was constant over time and that TCDD elimination followed a simple linear kinetic, that is, $elimination(t) = k_e A(t)$. After adjusting TCDD measurements for German background exposure levels the elimination rate constant k_e was estimated from workers with multiple TCDD measurements. TCDD working exposure levels in the working areas were estimated by modelling the $intake(t)$ -function. These estimates were used to compute individual time courses of TCDD exposure for all cohort members, which served as basis for the risk assessment, see Flesch-Janys et al. (1995) and Becher et al. (1998) for details.

Uncertainty in the backcalculation of the TCDD concentrations was investigated recently by Heinzl and Edler (2002) and Edler, Heinzl and Mittlboeck (2004). Two potentially crucial model assumptions for estimating the exposure of a person are the assumption of lifetime constancy of total lipid volume (TLV) of the human body and the assumption of a simple linear kinetic of TCDD elimination. In 1995 a modified Michaelis-Menten kinetic, also known as Carrier kinetic, has been suggested to link the TCDD elimination rate to the available TCDD amount in the body (Carrier et al. 1995a, 1995b). That is, TCDD elimination would be faster, of nearly the same rate, or slower under this kinetic than under a simple linear kinetic when the individual would be highly, moderately, or just slightly contaminated, respectively.

A Monte Carlo simulation study was designed in order to mimick the essential features of the Boehringer cohort's exposure history (Heinzl and Mittlboeck 2004). Five main working areas with lognormally distributed individual TCDD exposure were assumed with mean intakes of 3500, 150, 40, 5 and 0 ng_{TCDD}/kg_{fat}/year, respectively. The highest exposure of 3500 ng_{TCDD}/kg_{fat}/year mean intake was assumed as occurred only in the 1950ies. Stop of exposure was assumed at plant closure in 1984 and determination of TCDD concentrations in cohort

members was simulated to happen in the early 1990ies, when most measurements in this cohort were performed. Mean background exposure was set to 1 ng_{TCDD}/kg_{fat}/year. Hiring, change of working area, termination of contract, retirement and death of the workers were simulated in the Monte Carlo study. Each worker was assumed as of being under permanent hazard to develop cancer. The increase in cancer hazard was linked to TCDD exposure via the AUC dose metric. For a virtual worker diseased of cancer an increased mortality hazard was assumed and – provided still active – retirement was entailed.

The simulation study consisted of three main steps: data generation, exposure back-calculation and dose-response modelling. Data were generated according to four different scenarios (see Table 1). The age-varying TLV values (scenarios II-IV) were randomly generated by adapting formulas and results reported by Thomaseth and Salvan (1998). These authors also provided the elimination function for scenario III. The elimination function of scenario IV is the Carrier kinetic (Carrier et al. 1995a, 1995b). Exposure back-calculation of the simulated data was always performed as it had been in the original statistical analyses of the Boehringer cohort, that is, constant TLV over time and simple linear elimination kinetic were assumed. For modelling the dose-response relationship of TCDD exposure on cancer mortality time-dependent AUC values were used as explanatory variable in a Cox proportional hazards regression model (Cox 1972, Flesch-Janys et al. 1995, Becher et al. 1998). The simulation scenarios I, II and III were replicated 100 times each. The elimination function of scenario IV resulted in a computationally time-consuming procedure such that only 30 replications were performed. The SAS software system was used for Monte Carlo simulations and statistical computations (SAS Institute Inc., Cary, NC, USA).

2.3 Results

Figure 2a shows the simulation results of working area 1 (highest exposure). This working area only existed in the 1950ies with a mean exposure of 3500 ng_{TCDD}/kg_{fat}/year. Figures 2b

and 2c show the results for working areas 2 and 3 of medium level exposures of 150 and 40 ng_{TCDD}/kg_{fat}/year, respectively. Figure 2d exhibits the results for working area 4 with a mean exposure of 5 ng_{TCDD}/kg_{fat}/year, which is just above background exposure. The results obtained for the non-exposed working area 5 are shown in Figure 2e.

The main consequences of the results of the simulation study for the exposure backcalculation method are evident. It is no surprise that a back-calculation procedure based on the assumptions of scenario I performs best for data generated under scenario I. If scenarios II or III would be true, however, then a back-calculation procedure based on the assumptions of scenario I would yield biased exposure estimates, but this bias would not be too bad throughout. In contrast, if a Carrier elimination kinetic (as in scenario IV) would be true, then high levels of TCDD exposure would be underestimated and low levels would be overestimated, respectively, when exposure back-calculation was based on the assumptions of scenario I. In other words, the TCDD dose metric is compressed if a linear kinetic is used to reconstruct an actually underlying Carrier kinetic.

The impact of carrying-over the uncertainty in TCDD-dose estimates onto cancer response estimates are shown in Figures 3a and 3b. The estimated Cox regression coefficients under scenarios I- III are virtually unbiased regardless of the existence of a TCDD effect on cancer mortality. Under scenario IV (the Carrier kinetic) the variance of the estimated coefficients increases and in case of a TCDD effect a considerable bias is observed as well. Both effects are natural consequences of the TCDD dose compression mentioned above. That is, if exposure is estimated by assuming a linear elimination kinetic although a Carrier kinetic actually holds, high exposures in reality will be underestimated through statistical analysis and low exposures will be overestimated, respectively. This inevitably affects the resulting individual concentration-time curves and the derived TCDD dose metric values.

2.4 Discussion

The results of the simulation study directly lead to the question how plausible the Carrier kinetic is from an empirical viewpoint. A literature search e.g. reveals that repeated TCDD measurements of two highly exposed Viennese women supports the Carrier kinetic (Geusau et al. 2002) as well as a combined evaluation of the Seveso and Ranch Hand data sets also give some evidence in this direction (Michalek et al. 2002). Is this enough evidence or plausibility to use the Carrier kinetic in the future? Who can/should decide this question? And even if common agreement could be reached that the Carrier kinetic is to be preferred, then still the question would remain: "How much can we be wrong and still be fine?" This immediately transforms into the question how much is "how much"? Again, who can/should decide?

Another type of question concerns the precision of an uncertainty assessment. Obviously, uncertainty assessment comprises uncertainty itself as during an uncertainty assessment various assumptions have to be made. Actually, this would require an uncertainty assessment of its own. That is, the idea of a meta-uncertainty may be put forward - the uncertainty of the uncertainty assessment. Consequently, there is as well meta-meta-uncertainty, the uncertainty of the meta-uncertainty assessment such that one layer of uncertainty can be built on another. Obviously, this rather automatic approach of dealing with impreciseness of an uncertainty assessment would end up in a quandary (Heinzl and Edler 2003). What can/should be done instead?

3. The role of *ex post* uncertainty assessment

This Section aims to provide some answers to the questions raised in the previous Section.

After the results of an *ex post* uncertainty assessment have been provided it is inevitable to consider proper consequences. Three issues arise: (a) the plausibility of the examined newly emerged aspects, (b) the quality of the uncertainty assessment, and (c) the acceptable amount

of uncertainty. For us these issues meet in the understanding that an uncertainty assessment has to be embedded in a broad scientific discussion process. Accordingly, results and consequences of an uncertainty assessment should be evaluated by an expert panel formed by statisticians, toxicologists, risk assessors, risk managers and other stakeholders from science and society. This panel should elaborate new aspects of the problem and it should formulate further questions to be addressed by the uncertainty assessment team.

In our opinion, a time schedule for the *ex post* uncertainty assessment, the above mentioned panel and a fund-raising scheme should be already set up during the original risk assessment, latest, however, at the time of the original risk management decision. At the lapse of time (and in the case of an unforeseen but vitally important event even before the lapse of time) that panel would have to decide on what new scientific aspects would justify an *ex post* uncertainty assessment. Subsequently, the uncertainty assessment team will be engaged. The panel will provide mandatory specifications and supervise and foster progress of work. It will critically evaluate intermediate results in order to flexibly calibrate the goals of the *ex post* uncertainty assessment. As such it will work like a steering committee of a large project.

After the availability of the uncertainty assessment team's results the panel will use them to re-evaluate the original risk assessment. Three main conclusions are possible. Firstly, the panel decides that the original risk assessment is still based on solid scientific ground. Secondly, the *ex post* uncertainty assessment may confirm the original risk assessment but may raise some doubts concerning various details. As a consequence new empirical studies or further mechanistic research may be requested, e.g. in the TCDD-case the availability of new human data will be rather limited whereas it will pose no problem to perform animal studies either way. It is also imaginable that the panel may request repeating selected statistical analyses of the original risk assessment process, e.g. in the TCDD-case the detected uncertainty in the high-dose exposure levels may directly lead to a request for a re-evaluation of the low dose extrapolation. Thirdly, the strongest consequence of the panels work could be

a request for re-evaluation of the entire original risk assessment.

Prerequisites for a successful *ex post* uncertainty assessment are the availability of a detailed and profound documentation of the statistical analysis in the original risk assessment and the availability of reasonably endowed funds.

The uncertainty assessment approach exemplarily presented in Section 2 is based on the investigation of „what happens if“-scenarios. A more refined approach is described in Vose (2000), and a fully Bayesian approach to uncertainty assessment is described in Bois and Diack (2005).

It is an open question whether members of the original risk assessment team should be part of the *ex post* uncertainty assessment team. Efficient work due to already existing knowledge is counterbalanced by some hardly avoidable factory blindness. A possible solution could be the engaging of multiple uncertainty assessment teams which work independently from each other. The members of the original risk assessment team may form one of these uncertainty assessment teams. Of course, there may be also other plausible reasons for engaging multiple uncertainty assessment teams. In the TCDD-case the Boehringer cohort is just one of several cohorts which provided evidence for the risk management decision. It seems then plausible to engage an uncertainty assessment team for each cohort.

Finally note that an *ex post* uncertainty assessment can only shed light onto overlooked issues or issues which have not been known at the time of original risk assessment itself. It is probable that some time after the completion of the *ex post* uncertainty analyses again new scientific aspects may evolve. In the case of the dioxins, for example, one may feel urged to include exposure to polychlorinated biphenyls (PCBs) into the risk assessment. Then further *ex post* uncertainty assessments should be performed.

References

- [1] Becher, H., Flesch-Janys, D., Gurn, P. and Steindorf, K. (1998). Krebsrisikoabschätzung für Dioxine [Quantitative cancer risk assessment for dioxins]. Berlin: Erich Schmidt Verlag.
- [2] Bois F.Y., Diack C.T. (2005). Uncertainty Analysis: the Bayesian approach. In: Quantitative Methods for Cancer and Human Health Risk Assessment (Eds. Edler L., Kitsos C.P.), Chichester: Wiley, to appear.
- [3] Carrier, G., Brunet, R. C. and Brodeur, J. (1995a). Modeling of the toxicokinetics of polychlorinated dibenzo-p-dioxins and dibenzofurans in mammals, including humans. I. Nonlinear distribution of PCDD/PCDF body burden between liver and adipose tissues. *Toxicology and Applied Pharmacology*, 131, 253-266.
- [4] Carrier, G., Brunet, R. C. and Brodeur, J. (1995b). Modeling of the toxicokinetics of polychlorinated dibenzo-p-dioxins and dibenzofurans in mammals, including humans. II. Kinetics of absorption and disposition of PCDD/PCDFs. *Toxicology and Applied Pharmacology*, 131, 267-276.
- [5] Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society B*, 34, 187-220.
- [6] Edler, L., Heinzl, H. and Mittlboeck, M. (2004). Carrying-over toxicokinetic model uncertainty into cancer risk estimates: the TCDD example. *Organohalogen Compounds*, 66, 3356-3362.
- [7] Flesch-Janys, D., Berger, J., Gurn, P., Manz, A., Nagel, S., Waltsgott, H. and Dwyer, J. H. (1995). Exposure to polychlorinated dioxins and furans (PCDD/F) and mortality in a cohort of workers from a herbicide-producing plant in Hamburg, Federal Republic of Germany. *American Journal of Epidemiology*, 142, 1165-1175. Published erratum appears in *American Journal of Epidemiology*, 144, 716 (1996).

- [8] Geusau, A., Schmaldienst, S., Derfler, K., Pöpke, O. and Abraham, K. (2002): Severe 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) intoxication: kinetics and trials to enhance elimination in two patients. *Arch Toxicol*, 76, 316-325.
- [9] Heinzl, H. and Edler, L. (2002). Assessing uncertainty in a toxicokinetic model for human lifetime exposure to TCDD. *Organohalogen Compounds*, 59, 355-358.
- [10] Heinzl, H. and Edler, L. (2003). Evaluating and assessing uncertainty of large-scaled statistical analyses exemplified at the Boehringer TCDD cohort. RM 2003, Proceedings of the Second Workshop on Research Methodology, Amsterdam: VU University, 87-94.
- [11] Heinzl, H. and Mittlboeck, M. (2004). Design aspects of a computer simulation study for assessing uncertainty in human lifetime toxicokinetic models. *Compstat 2004, Proceedings in Computational Statistics*, Heidelberg: Physica-Verlag, 199-211.
- [12] IARC (1997). IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol. 69: Polychlorinated Dibenzo-para-dioxins and Polychlorinated Dibenzofurans. International Agency for Research on Cancer, Lyon.
- [13] McGregor, D.B., Partensky, C., Wilbourn, J. and Rice J.M. (1998). An IARC Evaluation of Polychlorinated Dibenzo-p-dioxins and Polychlorinated Dibenzofurans as Risk Factors in Human Carcinogenesis. *Environ Health Perspect* 106 (Suppl 2), 755-760.
- [14] Michalek, J. E., Pirkle, J. L., Needham L. L., Patterson Jr., D. G., Caudill, S. P., Tripathi, R. C. and Mocarelli P. (2002). Pharmacokinetics of 2,3,7,8-tetrachlorodibenzo-p-dioxin in Seveso adults and veterans of operation Ranch Hand. *Journal of Exposure Analysis and Environmental Epidemiology*, 12, 44-53.
- [15] Thomaseth, K. and Salvan, A. (1998). Estimation of occupational exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin using a minimal physiologic toxicokinetic model.

Environmental Health Perspectives, 106 (Supplement 2), 743-753. Published erratum appears in Environmental Health Perspectives, 106 (Supplement 4), CP2 (1998).

[16] Vose, D. (2000). Risk analysis: a quantitative guide. Second edition, Chichester: Wiley.

Scenario	total lipid volume (TLV)	elimination function
I	constant over time	simple linear kinetic
II	varying with worker's age	simple linear kinetic
III	varying with worker's age	simple linear kinetic with modifications according to Thomaseth and Salvan (1998)
IV	varying with worker's age	Carrier kinetic (Carrier et al. 1995a, 1995b)

Table 1: Data generation scenarios.

Figure legends

Figure 1:

Schematic representation of the situation after the risk management decision.

Figures 2a, 2b, 2c, 2d and 2e:

Simulation results of estimated TCDD exposure in working area 1 (extremely high exposure only in the 1950ies), working areas 2 and 3 (medium level exposure), working area 4 (low exposure) and working area 5 (no exposure). The horizontal lines denote the true mean exposure levels.

Figures 3a and 3b:

Simulation results of relating the dose estimates derived from back-calculation to cancer mortality by employing a Cox proportional hazards regression model. No and small TCDD effects were assumed in the simulations. Positive/negative regression coefficients stand for an increase/decrease in mortality hazards, respectively. The horizontal lines at the value of zero stand for the no TCDD effect situation which is the truth in Figure 3a. In Figure 3b the truth lies around the median of scenario I.

Figure 1:

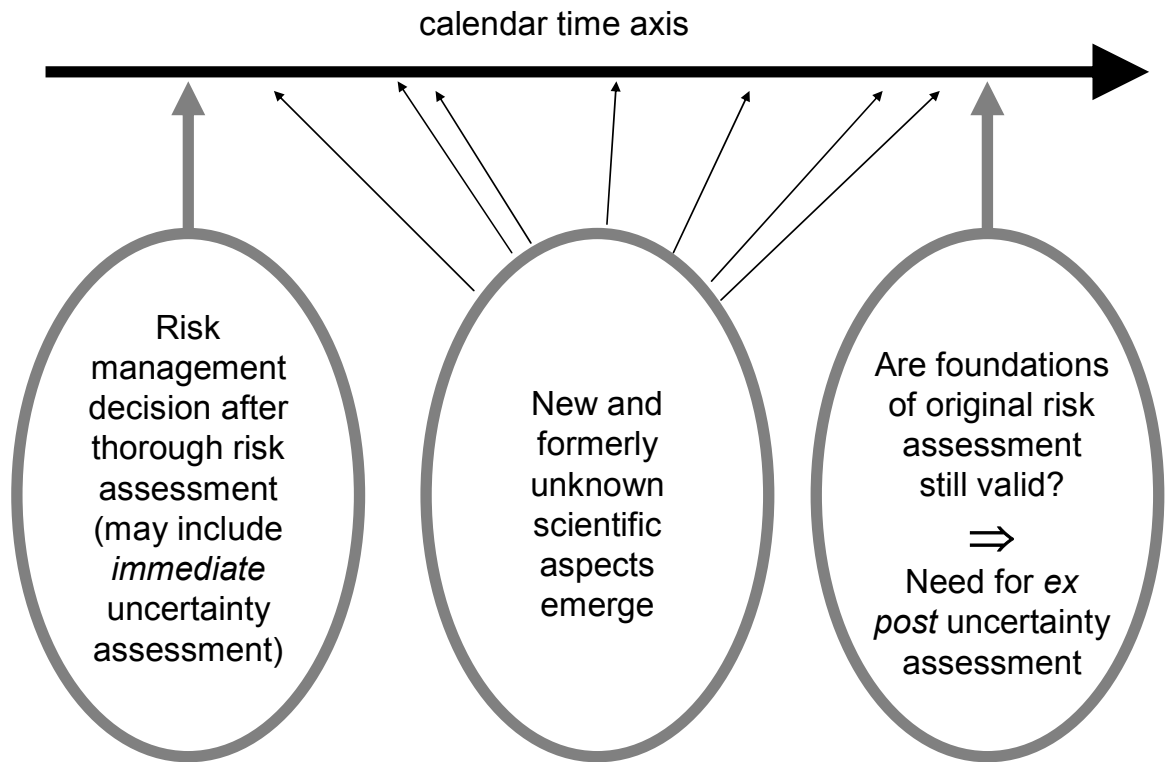
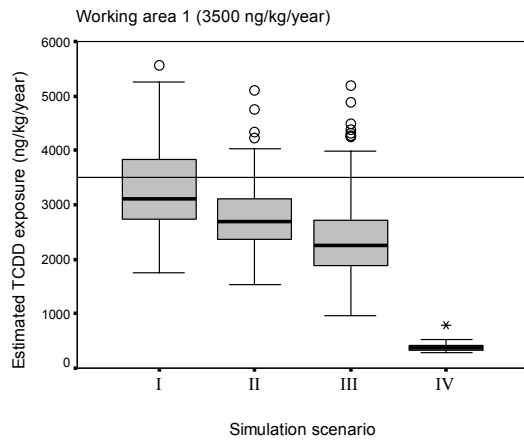


Figure 2a:



Figures 2b:

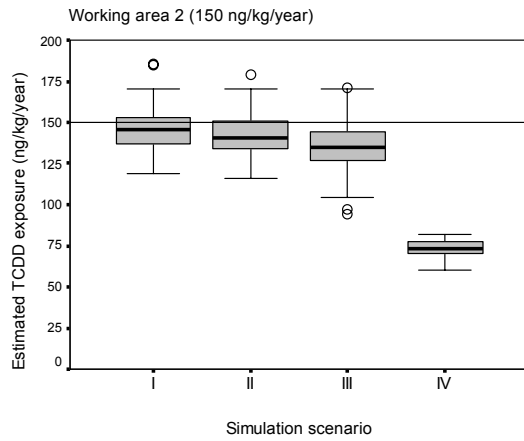
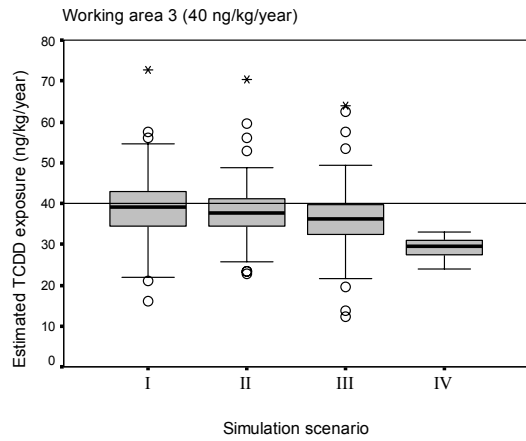


Figure 2c:



Figures 2d:

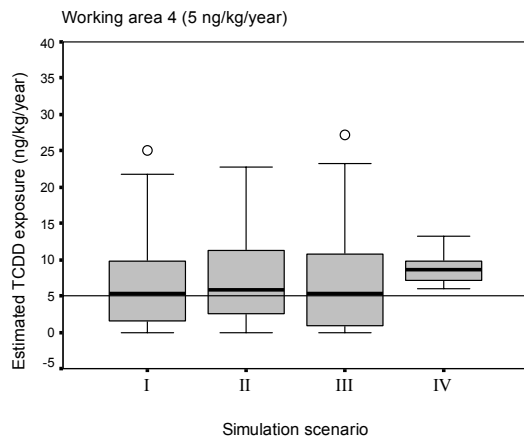
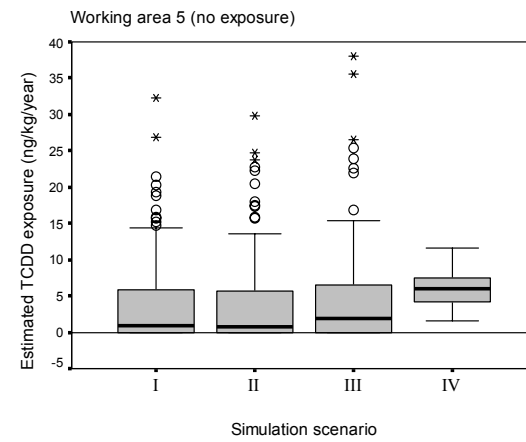


Figure 2e:



Figures 3a:

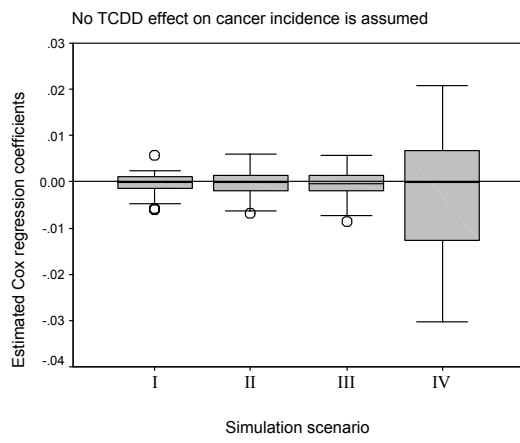


Figure 3b:

