Population toxicokinetics of ethylene: Models and validation of first order assumptions on kinetic processes

Silvia Selinski[1], Barbara Quinke[2], Klaus Golka[2],
Hermann M. Bolt[2], Wolfgang Urfer[1]

[1]Fachbereich Statistik, Universität Dortmund, 44221 Dortmund
[2]Institut für Arbeitsphysiologie an der Universität Dortmund, Ardeystraße 67, 44139 Dortmund

Abstract

The determination of toxicokinetic parameters is an essential component in the risk assessment of potential harmful chemicals. It is a key step to analyse the processes involved in the formation of DNA adducts which are connected with the development of chemical-induced cancer.

A general problem is the extrapolation of toxicological data from experimental animals to the human organism. Therefore a valid characterisation of the relevant processes for the whole species is required, i.e., of population mean parameters instead of sets of parameters for different individuals. These, again, may vary between repeated experiments at the same or at different administered doses. Nevertheless, these differences are of great importance in obtaining a more precise insight into the variability structure of process investigated within the test animal population, so that a valid basis for further research is the final result.

The theory of hierarchical models, particularly the work of Racine-Poon (1985) and Racine-Poon and Smith (1990), provides a procedure which incorporates both, modelling of the variability structure and estimation of population mean parameter vectors. The present study was designed to elucidate interindividual and interoccasion variability of toxicokinetic parameters relevant for the biological transformation of one of the basic petrochemical industrial compounds, ethylene (ethene), which is also a physiological body constituent, to its metabolite, ethylene oxide, which is a proven carcinogen.

Key words: Ethylene, ethylene oxide, toxicokinetics, population model, repeated measurements, EM algorithm, interindividual variability, interoccasion variability

1. Introduction

The determination of toxicokinetic parameters is an essential component in the risk assessment of potential harmful chemicals. Most chemical carcinogens are transformed into a chemical active form, its metabolite, that is able to interact with cellular macromolecules such as DNA, RNA, and protein, and might finally lead to the development of cancer. The relationship between applied dose and tumor response is nonlinear (Filser and Bolt, 1984). This non-linearity is supposed to be connected with the kinetic processes involved in the formation of DNA adducts (Hoel *et al.*, 1983). Hence an important step to assess the risk of a xenobiotic is to investigate the kinetic processes of its uptake, metabolism, and exhalation.

The recognition of wide genetic variations in human metabolism of foreign chemicals has focussed toxicological interests on pharmacogenetic factors in experimental toxicological studies (Lovell, 1993). In particular, the importance of the genetic make-up of a test animal population for experimental toxicity testing has been stressed in view of the wide genetically determined variability of toxicokinetic and toxicodynamic relationships in natural populations (Hedrich and Löscher, 1993). The experimental use of inbred strains of rodents is an important tool to reduce the biological interindividual variability of toxicological responses to chemicals, much in contrast to the situation in humans to which a toxicity extrapolation is made.

It is now established that much of the differences in toxicodynamic responses to chemicals within a population is based on matters of toxicokinetics, in particular on genetically imprinted differences in activities of enzymes involved in the metabolism of foreign compounds. It is therefore important to define the intrinsic experimental variability of toxicokinetic factors between the members of inbred rodent strains used for toxicological studies.

The present study has been designed to elucidate interindividual and interoccasion variability of toxicokinetic parameters relevant for the carcinogenicity of one of the basic petrochemical industrial compounds, ethylene (ethene).

A two-compartment model is used to describe the processes of uptake, exhalation, and metabolic elimination of ethylene approximating the real kinetic processes by first order kinetics. Two kinds of experimental designs are investigated: Repeated exposure to equal (group A) and to different (group B) doses. We apply nonlinear hierarchical models estimating the individual and population mean parameters as well as the unknown covariance matrices by the use of an EM algorithm. Furthermore, we provide a method to check the assumptions of first order kinetics.

2. Project and Data

The aim of this investigation is to determinate the population mean kinetic parameters of uptake, exhalation, and metabolism of the chemical ethylene and to quantify the variability due to interindividual and interoccasion differences where the conditions at repeated occasions may be equal (group A) or different (group B).

Ethylene is one of the basic petrochemical industrial compounds. In the living mammalian organism, ethylene is partly transformed, by hepatic metabolising enzymes (cytochrome P-450) to ethylene oxide (Filser and Bolt, 1983) which is biologically reactive and thereby genotoxic (Kirkovski *et al.*, 1998). The principles of the toxicokinetics of this transformation have been extensively studied (Filser and Bolt, 1984; Bolt *et al.*, 1984), and estimates of the carcinogenic risk of ethylene based on its metabolic transformation to ethylene oxide were published (Filser and Bolt, 1984; Thier and Bolt, 2000). Recent interest has been focussed on "endogenous" carcinogenic risks of ethylene. Ethylene is not only an exogenous and potentially toxic foreign chemical, but also a physiological body constituent (Filser *et al.*, 1992; Bolt, 1996; Bolt *et al.*, 1997). This particular aspect has a potential impact for legal regulations of weak genotoxins in general (Filser *et al.*, 1994; Bolt, 1998)

Previous inhalation experiments with ethylene have indicated that the metabolism may be well approximated by first order kinetics at concentrations below 800 *ppm* (*parts per million)*. At higher concentrations the metabolism of ethylene becomes more and more saturated (Bolt and Filser, 1987).

Experimental design

Two different groups of experiments were investigated at the *Institute of Occupational Physiology at the University of Dortmund*, each with 10 male Sprague-Dawley rats. The animals had an average weight of 300 *g*.

Both groups of experiments were carried out using the "closed chamber technique" as reviewed by Filser (1992), which allows investigations of kinetics of volatile chemicals *in vivo* (cf. Quinke *et al.*, 2000; Selinski *et al*., 2000, for further details).

The experiments of the first group (group A) had the following design:

Each of the ten rats was exposed to an initial concentration of about 100 *ppm* ethylene for a time period of about 8 hours. In that time about 20 samples per animal were taken, i.e., one sample every 25 minutes. This procedure was repeated four times with the same initial concentration of about 100 *ppm* ethylene, so that we finally received five experimental series per animal observed under identical conditions.

The design of the second group of experiments (group B) differed in so far from the previous as each of the ten further rats was observed at different and increasing initial concentrations of 20, 50, 100, 200 and 500 *ppm* ethylene (cf. Quinke *et al*., 2000; Selinski and Urfer, 1998, for further details).

Note, that due to the experimental conditions the initial concentrations are not exactly known and have to be treated as additional parameters.

3. Models and Methods

The following section presents the toxicological models as basis of the data analysis, a method for checking an essential part of these assumptions – overall first order kinetics and introduces the statistical models and the computational formulas.

3.1 Two-compartment model

The two-compartment model used by Filser (1992) for the characterisation of exposure to volatile xenobiotics describes uptake, endogenous production, excretion, and the metabolic elimination of the substance. The model is depicted as follows: a xenobiotic gas, in this case ethylene, enters the body and is exhaled. This process is described by introducing two compartments, the first, C_1, representing the environment outside the body, here the inhalation chamber of the exposition system, and the second compartment, C_2, the body itself. The volatile xenobiotic migrates from one compartment to the other through a theoretical interface. During this process, some portion of the xenobiotic within the organism, at any stage, is eliminated by metabolic processes, and another portion is again exhaled (cf. Fig. 1).

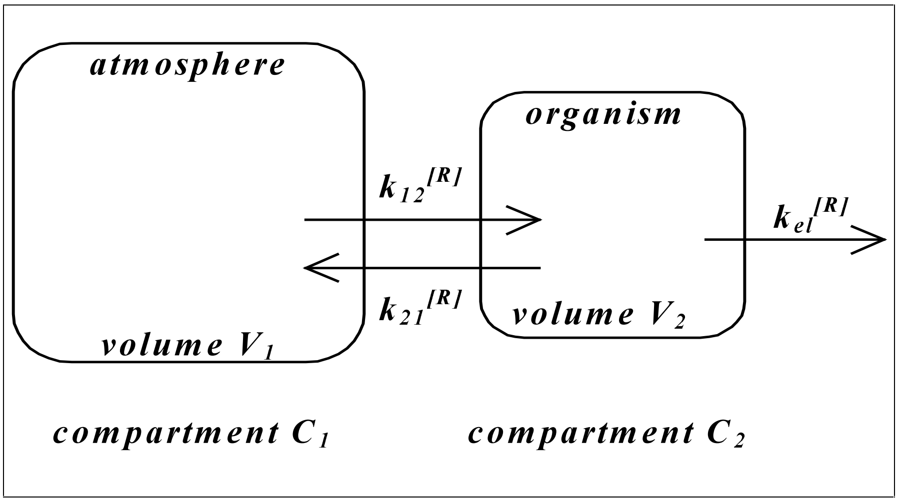

Figure 1. *Two-compartment block model in the case of metabolic turnover*

3.1.1 First order kinetics

This paper concentrates on overall first order kinetic processes which seem to be a valid approximation of the true processes within the applied range of concentrations of 20 to 500 *ppm* ethylene (Bolt and Filser, 1987; Selinski *et al.*, 2000).

Let $y_l(t)$, $l = 1, 2$, denote the concentration of a xenobiotic in compartment l at time t and let V_l describe the volume of the compartment.

In the case of overall first order kinetics, each partial process can be characterised by one rate or velocity constant k, that is $k_{12}^{[R]}$ for the uptake, $k_{21}^{[R]}$ for the exhalation, and

$k_{el}^{[R]}$ for the metabolic elimination (cf. Fig. 1). Thus the two-compartment model can be described as follows (Becka *et al.*, 1993):

The concentration in the first compartment (atmosphere) is given by

$$y_1(t) = y(0) \cdot \left\{ \frac{\left(k_{12}^{[R]} + \lambda_1\right)\exp\{\lambda_2 t\} - \left(k_{21}^{[R]} + \lambda_2\right)\exp\{\lambda_1 t\}}{(\lambda_1 - \lambda_2)} \right\}, \tag{1}$$

the concentration in the second compartment (organism) is given by

$$y_2(t) = y(0) \cdot \left\{ \frac{\left(k_{12}^{[R]} + \lambda_1\right)\left(k_{12}^{[R]} + \lambda_2\right)}{(\lambda_1 - \lambda_2)\alpha_2 k_{21}^{[R]}} \cdot \left[\exp\{\lambda_2 t\} - \exp\{\lambda_1 t\}\right] \right\} \tag{2}$$

where $\lambda_{1,2} = \frac{1}{2}\left\{ -\left(k_{12}^{[R]} + k_{21}^{[R]} + k_{el}^{[R]}\right) \pm \sqrt{\left(k_{12}^{[R]} + k_{21}^{[R]} + k_{el}^{[R]}\right)^2 - 4k_{12}^{[R]}k_{el}^{[R]}} \right\}$, $\alpha_2 := V_2/V_1$ is the ratio of volumes, and $y(0)$ is the initial concentration in compartment 1 (Urfer and Becka, 1996).

3.1.2 Standardisation

In the practical application we have to take into account, that the individual organisms have different volumes which are also varying between repeated experimental occasions. According to Filser (1992) the individual rates of uptake $k_{12}^{[R]}$, exhalation $k_{21}^{[R]}$ and metabolic elimination $k_{el}^{[R]}$ are related to the respective rates k_{12}, k_{21} and k_{el} for a standard rat of 1000 *ml* by

$$\begin{aligned} k_{12}^{[R]} &= k_{12} \cdot v_2^{2/3}, && \\ k_{21}^{[R]} &= k_{21} \cdot v_2^{1/3}, && \text{and} \\ k_{el}^{[R]} &= k_{el} \cdot v_2, && \text{where} \end{aligned} \tag{3}$$

$v_2 = \left(\frac{1000}{V_2}\right)$ depends on the actual volume of the organism V_2 and the standard volume 1000 *ml*.

Substituting the real kinetic parameters in the respective formulas yields

$$f(\beta, t) = y_1(t) = y(0) \cdot \left\{ \frac{\left(k_{12}v_2^{2/3} + \lambda_1\right)\exp\{\lambda_2 t\} - \left(k_{21}v_2^{1/3} + \lambda_2\right)\exp\{\lambda_1 t\}}{(\lambda_1 - \lambda_2)} \right\}, \tag{4}$$

and

$$y_2(t) = y(0) \cdot \left\{ \frac{\left(k_{12}v_2^{2/3} + \lambda_1\right)\left(k_{12}v_2^{2/3} + \lambda_2\right)}{(\lambda_1 - \lambda_2)\alpha_2 k_{21}v_2^{1/3}} \cdot \left[\exp\{\lambda_2 t\} - \exp\{\lambda_1 t\}\right] \right\}, \text{ where} \tag{5}$$

$$\lambda_{1,2} = \frac{1}{2}\left\{-\left(k_{12}v_2^{2/3} + k_{21}v_2^{1/3} + k_{el}v_2\right) \pm \sqrt{\left(k_{12}v_2^{2/3} + k_{21}v_2^{1/3} + k_{el}v_2\right)^2 - 4k_{12}k_{el}v_2^{5/3}}\right\},$$

$\beta = (k_{12},\ k_{21},\ k_{el},\ y(0))^T = (\varphi^T,\ y(0))^T$ is the vector of the standardised kinetic parameters $\varphi = (k_{12}, k_{21}, k_{el})^T$ and the initial concentration $y(0)$.

3.2 Hierarchical models for repeated application of equal (group A) or different (group B) doses

3.2.1 Notation

The observed concentrations of ethylene in the atmosphere of the exposition system (compartment 1) are denoted by y_{ijk}, with

$i = 1, \ldots, 10$ (group A) or rather $i = 11, \ldots, 20$ (group B) the number of the individual rat

$j = 1, \ldots, J$ the index of the time point t_j and

$k = 1, \ldots, 5$ the number of the experiments.

The functional relationship is given by

$$y_{ijk} = f(\beta_{ik}, t_j) + \varepsilon_{ijk},$$

$i = 1, \ldots, 10$ or rather $i = 11, \ldots, 20$, $j = 1, \ldots, J$, $k = 1, \ldots, 5$,

where $f(\beta_{ik}, t_j)$ is a non-linear function of the individual parameter vector β_{ik} and the time t. The function $f(\beta_{ik}, t_j)$ denotes the expected concentration-time curve of the ith individual at the kth occasion.

In the present application the function f is derived from the two-compartment model and is given by eq. (4) substituting β by β_{ik}, so that

$$f\left(\beta_{ik}, t_j\right) = y_{ik}(0) \cdot \left\{\frac{\left(k_{12ik}v_{2ik}^{2/3} + \lambda_{1ik}\right)\exp\{\lambda_{2ik}t_j\} - \left(k_{21ik}v_{2ik}^{1/3} + \lambda_{2ik}\right)\exp\{\lambda_{1ik}t_j\}}{\left(\lambda_{1ik} - \lambda_{2ik}\right)}\right\}, \quad (6)$$

where $v_{2ik} = \left(\frac{V_{2ik}}{1000}\right)$ depends on the volume of the ith rat at the kth occasion V_{2ik} and

$$\lambda_{1ik,2ik} = \frac{1}{2}\left\{-\left(k_{12ik}v_{2ik}^{2/3} + k_{21ik}v_{2ik}^{1/3} + k_{elik}v_{2ik}\right) \pm \sqrt{\left(k_{12ik}v_{2ik}^{2/3} + k_{21ik}v_{2ik}^{1/3} + k_{elik}v_{2ik}\right)^2 - 4k_{12ik}k_{elik}v_{2ik}^{5/3}}\right\}$$

with $\lambda_{2ik} < \lambda_{1ik} < 0$.

The parameter vector $\beta_{ik} = (k_{12ik}, k_{21ik}, k_{elik}, y_{ik}(0))^T = (\varphi_{ik}^T, y_{ik}(0))^T$, where $\varphi_{ik} = (k_{12ik}, k_{21ik}, k_{elik})^T$ represents the vector of the standardised kinetic parameters, differs from individual to individual and is of dimension $p = 4$.

Our main interest are not the individual responses to the experimental conditions but is focussed on a population mean process, which underlies the different individual processes. The individual parameter vectors φ_{ik} may be regarded as to vary at random across an individual mean parameter vector φ_i, which describes the general behaviour of the respective processes for that individual. Furthermore the individual mean processes are supposed to vary across a population mean process with parameter vector φ in the manner of a random sample. Additionally we suppose that the variances of the observed concentration-time curves differ from individual to individual and from occasion to occasion.

3.2.2 Hierarchical model in case of repeated application of equal doses (group A)

A Bayesian approach according to Racine-Poon (1985) and Racine-Poon and Smith (1990) is applied to the data. We are interested especially in the variation of the individual responses at different dosing occasions, the so called *interoccasion* variability, and the variation between the individuals, the *intersubject* variability.

Non-linear hierarchical model

We propose a four-stage non-linear hierarchical model assuming that our observations y_{ijk} of the concentration of ethylene in the atmosphere of the exposition system are independent and have the following distribution:

given β_{ik}, τ_{ik}^2: $\quad y_{ijk} \sim N(f(\beta_{ik}, t_j), \tau_{ik}^2) \quad i = 1, \ldots, 10, j = 1, \ldots, J$ and $k = 1, \ldots, 5$,

with $\quad \beta_{ik} = (\varphi_{ik}, y_{ik}(0))^T$, and $\varphi_{ik} = (k_{12ik}, k_{21ik}, k_{elik})^T$

given β_i, Ω_i: $\quad \beta_{ik} \sim N(\beta_i, \Omega_i)\, i = 1, \ldots, 10$ and $k = 1, \ldots, 5$,

with $\quad \beta_i = (\varphi_i^T, y_i(0))^T$, and $\varphi_i = (k_{12i}, k_{21i}, k_{eli})^T$,

given β, Σ: $\quad \beta_i \sim N(\beta, \Sigma) \quad i = 1, \ldots, 10$, with $\beta = (\varphi^T, y(0))^T$, and $\varphi = (k_{12}, k_{21}, k_{el})^T$

$p(\beta) \propto 1 \qquad \forall\, \beta \in IR^4$.

In case of the present application f is specified by eq. (6).

Linear hierarchical model

We obtain the Bayes estimates for the population mean and individual parameter vectors β, β_i and β_{ik} by transforming the non-linear hierarchical model into a linear one, such as provided by Lindley and Smith (1972). For that purpose the observations y_{ijk} are replaced by an "almost" sufficient statistic ζ_{ik} with

$$\zeta_{ik} \sim N(\beta_{ik}, \tau_i^2 C_{ik}), \qquad i = 1, \ldots, 10,\ k = 1, \ldots, 5.$$

In the case of uninformative priors for the variances τ_i^2 the maximum likelihood estimate of β_{ik} can be used as a good approximation for ζ_{ik} (Racine-Poon, 1985; Selinski, 2001).

The resulting linear hierarchical model is given by:

given β_{ik}, τ_i^2: $\quad \zeta_{ik} \sim N(\beta_{ik}, \tau_i^2 C_{ik}), \quad i = 1, \ldots, 10,\ k = 1, \ldots, 5$

given β_i, Ω_i: $\quad \beta_{ik} \sim N(\beta_i, \Omega_i), \quad i = 1, \ldots, 10,\ k = 1, \ldots, 5$

given β, Σ: $\quad \beta_i \sim N(\beta, \Sigma), \quad i = 1, \ldots, 10$

$p(\beta) \propto 1, \quad \forall \beta \in IR^4$.

where $\tau_{ik}^{-2} C_{ik}^{-1}$ is the information matrix (cf. Selinski and Urfer, 1998; Selinski, 2001, for computational formulas):

$$\tau_{ik}^{-2} C_{ik}^{-1} = E\left[-\frac{\partial^2}{\partial \beta_{ik} \partial \beta_{ik}^T} \ln L\left(y_{1,1,1}, \ldots, y_{10,J,5} \middle| \beta_{1,1}, \ldots, \beta_{10,5}, \tau_{1,1}^2, \ldots, \tau_{10,5}^2 \right) \right] \tag{7}$$

In case of known variances τ_{ik}^2, and covariance matrices Ω_i and Σ the Bayes estimates can be computed as expectations of the posterior distributions of β, β_i, and β_{ik}, $i = 1, \ldots, 10$, $k = 1, \ldots, 5$, which can be derived easily from the linear hierarchical model (Lindley and Smith, 1972; Selinski *et al.*, 2000). However, we have only vague knowledge about these covariance matrices, and the aim of our investigation is to gain information about just these covariances, especially with regard to the interoccasion and interindividual variability. Hence, we need a method to estimate both the parameter vectors and the covariance matrices. Such a method is presented in the following section.

Estimators in the case of unknown covariance matrices

In the case of unknown variances τ_{ik}^2, $i = 1, \ldots, 10$, $k = 1, \ldots, 5$, Racine-Poon and Smith (1990) suggest to replace them by suitable estimates $\hat{\tau}_{ik}^2$. Thus, we approximate the Bayes estimate of τ_{ik}^2 by (Selinski, 2000)

$$\hat{\tau}_{ik}^2 = \frac{1}{J} \cdot \sum_{j=1}^{J} \left(y_{ijk} - f\left(\zeta_{ik}, t_j \right) \right)^2 \qquad , i = 1, \ldots, 10, k = 1, \ldots, 5. \tag{8}$$

For the joined estimation of the individual and the population mean parameters as well as the covariance matrices $\Omega_1, \ldots, \Omega_{10}$ and Σ an EM-type iterative algorithm as proposed by Dempster *et al.* (1977) is adapted to our four stage model. We assume, that the inverse covariance matrices Ω_i^{-1}, $i = 1, \ldots, 10$, and Σ^{-1} follow Wishart distributions with degrees of freedom ρ_1 and ρ_2 and matrices R_1 and R_2, respectively. Thus $R_1^{-1}/(\rho_1 - p-1)$ and $R_2^{-1}/(\rho_2 - p-1)$ play the role of prior estimates of Ω_i and Σ. Vague knowledge about the inverse covariance matrices $\Omega_1^{-1}, \ldots, \Omega_{10}^{-1}$, and Σ^{-1} can be expressed by choosing ρ_1 and ρ_2 as small as possible, i. e. $\rho_1 = \rho_2 = p = 4$. The choice of R_1 and R_2, respectively, seems to have little influence on the estimates (Racine-Poon, 1985).

Substituting $\hat{\tau}_{ik}^2$ for τ_{ik}^2, if necessary, we obtain the approximations of the Bayes estimates at the lth iteration of the EM-algorithm, $\beta_{ik}^{(l)}$, $\beta_i^{(l)}$, $\beta^{(l)}$, $\Omega_i^{(l)}$, and $\Sigma^{(l)}$, $i = 1, \ldots, 10$, $k = 1, \ldots, 5$, by computing the expectation of the posterior distribution of β, β_i, and β_{ik},

replacing the covariance matrices by their current approximations $\Omega_1^{(l-1)}, \ldots, \Omega_{10}^{(l-1)}$, and $\Sigma^{(l-1)}$, (E-Step) and calculating $\Omega_1^{(l)}, \ldots, \Omega_{10}^{(l)}$, and $\Sigma^{(l)}$ afterwards as the posterior modes using $\beta^{(l)}$, $\beta^{(l)}$, and $\beta^{(l)}$, $i = 1, \ldots, 10$, $k = 1, \ldots, 5$ (M-Step) (Selinski *et al.*, 2000).

E-Step

Approximating $\Omega_1, \ldots, \Omega_{10}, \Sigma$ by $\Omega_1^{(l-1)}, \ldots, \Omega_{10}^{(l-1)}$, and $\Sigma^{(l-1)}$ we obtain

$$\beta^{(l)} = \left[\sum_{i=1}^{10} \sum_{k=1}^{5} \left(\hat{\tau}_{ik}^2 C_{ik} + \Omega_i^{(l-1)} + \Sigma^{(l-1)} \right)^{-1} \right]^{-1} \cdot \sum_{i=1}^{10} \sum_{k=1}^{5} \left(\hat{\tau}_{ik}^2 C_{ik} + \Omega_i^{(l-1)} + \Sigma^{(l-1)} \right)^{-1} \zeta_{ik} \,. \quad (9)$$

Using the current estimate $\beta^{(l)}$ of β yields

$$\beta_i^{(l)} = \left[\left[\sum_{k=1}^{5} \left(\hat{\tau}_{ik}^2 C_{ik} + \Omega_i^{(l-1)} \right)^{-1} \right] + \Sigma^{(l-1)-1} \right]^{-1} \cdot \left[\left(\sum_{k=1}^{5} \left(\hat{\tau}_{ik}^2 C_{ik} + \Omega_i^{(l-1)} \right)^{-1} \cdot \zeta_{ik} \right) + \Sigma^{(l-1)-1} \cdot \beta^{(l)} \right] \quad (10)$$

and

$$\beta_{ik}^{(l)} = \left[\left(\hat{\tau}_{ik}^2 C_{ik} \right)^{-1} + \left(\Omega_i^{(l-1)} + \Sigma^{(l-1)} \right)^{-1} \right]^{-1} \cdot \left[\left(\hat{\tau}_{ik}^2 C_{ik} \right)^{-1} \cdot \zeta_{ik} + \left(\Omega_i^{(l-1)} + \Sigma^{(l-1)} \right)^{-1} \cdot \beta^{(l)} \right]. \quad (11)$$

M-Step

Setting β_{ik}, β_i and β, $i = 1, \ldots, 10$, $k = 1, \ldots, 5$, equal to their current values $\beta^{(l)}$, $\beta^{(l)}$, and $\beta^{(l)}$, the conditional posterior mode is given by

$$\Omega_i^{(l)} = \frac{R_1^{-1} + \sum_{k=1}^{5} \left(\beta_{ik}^{(l)} - \beta_i^{(l)} \right) \cdot \left(\beta_{ik}^{(l)} - \beta_i^{(l)} \right)^T}{5 + \rho_1 - p - 1}, \qquad i = 1, \ldots, 10, \text{ and} \quad (12)$$

$$\Sigma^{(l)} = \frac{R_2^{-1} + \sum_{i=1}^{10} \left(\beta_i^{(l)} - \beta^{(l)} \right) \left(\beta_i^{(l)} - \beta^{(l)} \right)^T}{10 + \rho_2 - p - 1} \quad (13)$$

Both steps are repeated until $\Omega_1^{(l)}, \ldots, \Omega_{10}^{(l)}$, and $\Sigma^{(l)}$ converge. Racine-Poon (1985) suggests as criterion for convergence, that the maximum change in the elements of the covariance matrices between successive iterations should be less than 0.001.

Reasonable starting values $\Omega_1^{(0)}, \ldots, \Omega_{10}^{(0)}$, and $\Sigma^{(0)}$ are given by

$$\Omega_i^{(0)} = \frac{R_1^{-1} + \sum_{k=1}^{5} \left(\zeta_{ik} - \overline{\zeta}_{i.} \right) \left(\zeta_{ik} - \overline{\zeta}_{i.} \right)^T}{5 + \rho_1 - p - 2}, \qquad i = 1, \ldots, 10, \text{ and}$$

$$\Sigma^{(0)} = \frac{R_2^{-1} + \sum_{i=1}^{10} \left(\overline{\zeta}_{i.} - \overline{\zeta}_{..} \right) \left(\overline{\zeta}_{i.} - \overline{\zeta}_{..} \right)^T}{10 + \rho_2 - p - 3},$$

where $\overline{\zeta}_{i.} = \frac{1}{5} \sum_{k=1}^{5} \zeta_{ik}$ and $\overline{\zeta}_{..} = \frac{1}{10} \sum_{i=1}^{10} \overline{\zeta}_{i.} = \frac{1}{50} \sum_{i=1}^{10} \sum_{k=1}^{5} \zeta_{ik}$.

3.2.3 Hierarchical model in case of repeated application of different doses (group B)

Analysing the experiments of group B it has to be taken into account that the initial concentration varies from occasion to occasion.

Non-linear hierarchical model

As we are merely interested in the kinetic parameter we ignore the potential dependence between their estimates and the initial concentration. Otherwise we would receive a more complex model which would be much more difficult to estimate as it was the case for model A (Selinski, 2001). Moreover, assuming overall first order kinetics implies this independence, although we have to verify this assumption, of course. A suitable test is presented in section 3.3.

Hence, we propose a four-stage non-linear hierarchical model assuming that our observations y_{ijk} of the concentration of ethylene in the atmosphere of the exposition system are independent and have the following distribution:

given $\varphi_{ik}, y_{ik}(0), \tau^2_{ik}$: $y_{ijk} \sim N(f(\varphi_{ik}, y_{ik}(0), t_j), \tau^2_{ik})$ $i = 11, \ldots, 20, j = 1, \ldots, J,$
$k = 1, \ldots, 5,$
with $\beta_{ik} = (\varphi^T_{ik}, y_{ik}(0))^T$, and $\varphi_{ik} = (k_{12ik}, k_{21ik}, k_{elik})^T$

given φ_i, Ω_i: $\varphi_{ik} \sim N(\varphi_i, \Omega_i)$, $i = 11, \ldots, 20, k = 1, \ldots, 5,$
with $\varphi_i = (k_{12i}, k_{21i}, k_{eli})^T$,

given φ, Σ: $\varphi_i \sim N(\varphi, \Sigma)$ $i = 11, \ldots, 20,$
with $\varphi = (k_{12}, k_{21}, k_{el})^T$

$p(\varphi) \propto 1$ $\forall\ \varphi \in IR^3$.

Linear hierarchical model

The non-linear hierarchical model is transformed into a linear one by substituting the observations y_{ijk} by the Maximum-Likelihood estimates ζ_{ik}. Thus, we receive the following linear model:

given θ, V: $\tilde{\zeta} \sim N(\theta, V)$,

where $\tilde{\zeta} = (\tilde{\zeta}_{11,1}, \ldots, \tilde{\zeta}_{20,5})^T$, $\tilde{\zeta}_{ik} = \left(\hat{k}_{12ik}, \hat{k}_{21ik}, \hat{k}_{elik}\right)^T$ are the three first components of the Maximum-Likelihood estimate ζ_{ik} of β_{ik}, $\theta = (\theta_{11}, \ldots, \theta_{20})^T$, $\theta_i = (\varphi_{i,1}, \ldots, \varphi_{i5})^T$, $V = diag\{(\tau^2_{11,1}\tilde{C}_{11,1}), \ldots, (\tau^2_{20,5}\tilde{C}_{20,5})\}$, and $\tau^2_{ik}\tilde{C}_{ik}$ denotes the left upper 3×3 matrix of the inverse of the Information matrix $\left(\tau^2_{ik}C_{ik}\right)^{-1}$.

given ψ, Ω: $\theta \sim N(Z_2\psi, \Omega)$,

where , $\theta = (\theta_{11}, \ldots, \theta_{20})^T$, $\theta_i = (\varphi_{i,1}, \ldots, \varphi_{i5})^T$, $\psi = (\varphi_{11}, \ldots, \varphi_{20})^T$

$$\Omega = \begin{pmatrix} \Omega_{11} & & & & & \\ & \ddots & & & 0 & \\ & & \Omega_{11} & & & \\ & & & \Omega_{12} & & \\ & 0 & & & \ddots & \\ & & & & & \Omega_{20} \end{pmatrix}, \text{ and } Z_2 = \begin{pmatrix} I & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ I & 0 & & & \\ 0 & I & 0 & & \vdots \\ \vdots & \vdots & \vdots & & \\ 0 & I & 0 & & \vdots \\ \vdots & 0 & \ddots & & \\ & & & \ddots & 0 \\ \vdots & \vdots & & 0 & I \\ & & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & I \end{pmatrix}$$ is a suitable

design matrix.

given φ, Λ: $\psi \sim N(Z_3\varphi, \Lambda)$,

where $\varphi = (k_{12}, k_{21}, k_{el})^T$, $\Lambda = diag\{\Sigma, \ldots, \Sigma\}$, and $Z_3 = (I_3, \ldots, I_3)^T$ is a suitable design matrix,

$p(\varphi) \propto 1$, $\forall \varphi \in IR^3$.

The Bayes estimates and the computational formulas of the EM algorithm are the same as in case of equal doses (group A). Note, that the dimension p of the parameter vectors is three instead of four.

3.3 Checking assumption of overall first order kinetics

An essential assumption of the present modelling approach is the assumption that all kinetic processes under investigation may be approximated well by first order or linear kinetics. The following section presents a method to check if this approximation is valid.

In the case of ethylene preceding experiments suggest that the processes of uptake, exhalation, and metabolism of ethylene may be well approximated by first order kinetics for concentrations below the point of saturation of about 800 *ppm* ethylene (Bolt, 1998). In the present inhalation study doses of about 20 – 500 *ppm* ethylene were used so that overall first order kinetics should provide a valid approximation of the real non-linear kinetic processes.

However, the assumption of first order kinetics has to be checked to avoid critical departures from linearity. For this purpose the experiments of group B, which provide information about the behaviour of the kinetic processes at different doses, serve as database for a test of first order kinetics. Moreover the results are counterchecked by an explorative analysis.

3.3.1 A distribution-free test for departures from first order kinetics

Assuming first order kinetics means that the processes are independent from the initial concentration. This assumption of independence can be used to test if the data may be well approximated by linear kinetics (Becka, 1994). So, the concentration-time curves depend on the initial concentration only through the factor $y_1(0)$.

In the case of overall first order kinetics the standardised observations

$$y_{ij^*k}^{[ST]} = \frac{y_{ij^*k}}{y_{ik}(0)}, \; i = 1, \ldots, I, \; k = 1, \ldots, K, \tag{14}$$

with j^* denoting the index of the time point t_{j^*}, are independent from the initial concentration $y_{ik}(0)$ in compartment 1. Testing the null hypothesis of independence from the initial concentration a time point t_{j^*} is chosen, where observations are available for all individuals $i = 1, \ldots, I$, and dosing occasions $k = 1, \ldots, K$. Usually a later time point t_{j^*} should be chosen as possible departures from first order kinetics would result in a clearer dependency of the standardised observations and the doses due to the duration of the partial processes.

As the initial concentration is often not exactly known in toxicokinetic experiments, $y_{ik}(0)$ has to be estimated. If a model is already fitted the estimate of $y_{ik}(0)$ may be used for standardising the observations. Otherwise, the first observation provides usually an adequate approximation for the initial concentration.

A further approach is to test for independence of the kinetic parameters from the initial concentration. Assuming first order kinetics this independence should hold for the estimates of the standardised individual and experiment specific kinetic parameters, as the shape of the concentration-time curves does not depend on the dose. The term 'standardised kinetic parameters' means here that the influence of the animal's volume is eliminated according to the procedure provided by Filser (1992).

Thus, the null hypothesis of independence of the standardised rates of uptake k_{12ik}, exhalation k_{21ik}, and metabolism $k_{\mathrm{el}ik}$, respectively, from the initial concentration is tested separately for each parameter.

Hence, it is possible to detect the partial processes, which are not approximated well by first order kinetics. Moreover, detecting such departures from linearity the direction of correlation between parameters and initial concentration is of substantial interest.

Usually, the sample sizes of toxicokinetic studies are quiet small so that non-parametrical tests will be the method of choice. Additionally, these methods are more robust against outliers.

Becka (1994) provides a procedure to test for departures from first order kinetics based on the Spearman rank correlation coefficient r_S. Following this attempt the Kendall test for independence is applied. This test is based on the Kendall correlation coefficient τ. The required properties of the data are the same for both statistics, which also contain the same amount of information about the sample. However, $|\tau| < |r_S|$ in almost every case. The advantage of τ is that its distribution converges faster against the normal distribution than the distribution of r_S does (Büning and Trenkler, 1994).

Kendall's correlation coefficient τ is defined as follows (Hollander and Wolfe, 1999):

Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a random sample from a continuous bivariate population, i.e. the n bivariate observations are mutually independent and identically distributed.

The Kendall population correlation coefficient is defined as

$$\tau = \frac{K}{(n-1)n/2}, \qquad (15)$$

where K is the Kendall statistic

$$K = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} Q((x_i, y_i), (x_j, y_j)) \text{ and} \qquad (16)$$

$$Q((a,b),(c,d)) = \begin{cases} 1, & if\,(d-b)(c-a) > 0 \\ -1, & if\,(d-b)(c-a) < 0 \end{cases}, \; 1 \le i < j \le n, \qquad (17)$$

is the sign statistic.

In case of many ties the modified correlation coefficient τ^B is given by

$$\tau^B = \frac{n_c - n_d}{\sqrt{(n-1)n/2 - T_X}\,\sqrt{(n-1)n/2 - U_Y}}, \text{ with} \qquad (18)$$

$$T_X = \sum_{i=1}^{g} (t_i - 1)t_i \,/\, 2 \text{ and } U_Y = \sum_{j=1}^{h} (u_j - 1)u_j \,/\, 2,$$

where t_i is the size of the tied X group i, g is the number of tied X groups, u_j is the size of the tied Y group j, and h is the number of tied Y groups.

A test of independence of X and Y based on Kendall's τ is given by the following definition.

3.3.2 Kendall's test for independence

Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a random sample from a continuous bivariate population. The Kendall test of independence is defined as a test of

H_0 : X and Y are *mutually* independent ⇔

$F_{X,Y}(x,y) \equiv F_X(x)F_Y(y)$, $\forall\ (x,y)$ pairs

versus

a. Two-Sided Test

H_1 : *X and Y* are correlated ⇔ $\tau \neq 0$

b. One-Sided Upper-Tail Test

H_2 : *X and Y* are positively correlated ⇔ $\tau > 0$

c. One-Sided Lower-Tail Test

H_3 : *X and Y* are negatively correlated ⇔ $\tau < 0$

at the α-level of significance if

reject H_0 if
a. $|K| \geq k_{\alpha/2}$ *Two-Sided Test*
b. $K \geq k_{\alpha}$ *One-Sided Upper-Tail Test*
c. $K \leq -k_{\alpha}$ *One-Sided Lower-Tail Test*

otherwise do not reject.

Critical values are given for example by Hollander and Wolfe (1999) for sample sizes up to 40 and by Neave and Worthington (1992) for sample sizes up to 50.

For large sample sizes n the standardised Kendall correlation statistic K^* has asymptotically a normal distribution and is given by

$$K^* = \frac{K - E_0(K)}{\sqrt{\mathrm{var}_0(K)}} \tag{19}$$

where $E_0(K) = 0$ is the expected value of K under H_0 and

$$\mathrm{var}_0(K) = \frac{n(n-1)(2n+5)}{18} \tag{20}$$

is the null variance of K (Hollander and Wolfe, 1999).

When H_0 is true and n tends to infinity, K^* has an asymptotic $N(0, 1)$ distribution. Thus,

reject H_0 if
a. $|K^*| \geq z_{\alpha/2}$ *Two-Sided Test* (21)
b. $K^* \geq z_{\alpha}$ *One-Sided Upper-Tail Test* (22)
c. $K^* \leq -z_{\alpha}$ *One-Sided Lower-Tail Test* (23)

otherwise do not reject.

In the case of ties among the X and/or among the Y observations, replace the function $Q((a, b), (c, d))$ in (17) by

$$Q^*((a,b),(c,d)) = \begin{cases} 1 & if\,(d-b)(c-a) > 0 \\ 0 & if\,(d-b)(c-a) = 0 \\ -1 & if\,(d-b)(c-a) < 0 \end{cases} \tag{24}$$

and compute K with these modified paired sign statistics. Note, that the test is now only approximately, and not exactly, of significance level α.

Applying the large-sample approximation (19) it has to be taken into account, that the tied observations result in a reduced variability, while the expectation of K under H_0 is not affected. In the case of tied X and / or Y observations the null variance is given by

$$\begin{aligned}\operatorname{var}_0(K) = {} & \frac{\left\{n(n-1)(2n+5) - \sum_{i=1}^{g} t_i(t_i-1)(2t_i+5) - \sum_{j=1}^{h} u_j(u_j-1)(2u_j+5)\right\}}{18} \\ & + \frac{\left\{\sum_{i=1}^{g} t_i(t_i-1)(t_i-2)\right\}\left\{\sum_{j=1}^{h} u_j(u_j-1)(u_j-2)\right\}}{9n(n-1)(n-2)} \\ & + \frac{\left\{\sum_{i=1}^{g} t_i(t_i-1)\right\}\left\{\sum_{j=1}^{h} u_j(u_j-1)\right\}}{2n(n-1)}\end{aligned} \tag{25}$$

where g denotes the number of tied X groups, h the number of tied Y groups, t_i is the size of the tied X group i and u_j is the size of the tied Y group j.

If neither the X group nor the Y group contains tied observations, we have $g = h = n$ and $t_i = u_j = 1$. In that case each term involving $(t_i - 1)$ and / or $(u_j - 1)$ reduces to zero and (25) is equal to the usual null variance of K as given in equation (20).

Thus, in the case of large-sample size n and tied observations, compute K^* with the modified paired sign statistic (24) using the null variance of K as given by (25):

$$K^* = \frac{K}{\sqrt{\operatorname{var}_0(K)}} \tag{26}$$

Hence, the approximations (21), (22) or (23) can be applied.

3.3.3 Graphical analysis

As in toxicological studies the database for testing hypothesis about the underlying kinetic processes is usually sparse Becka (1994) suggests to countercheck the results graphically. For this purpose the estimates of the standardised kinetic parameters are compared with the maximum concentration in the respective compartments. Thus, the maximum likelihood estimates $\hat{k}_{12ik}$ and the Bayes estimates k^*_{12ik} of the rate of uptake are plotted against the estimated initial concentration $\hat{y}_{ik}(0)$ in the first compartment. Further, the maximum likelihood estimates $\hat{k}_{21ik}$ and $\hat{k}_{elik}$ as well as the Bayes estimates k^*_{21ik} and k^*_{elik} are plotted against the estimated maximum concentration in the second compartment at time point $t_{\max} := \frac{\ln(\lambda_1/\lambda_2)}{(\lambda_2 - \lambda_1)}$ (Selinski, 2001).

4. Results

The methods presented in the previous chapter were applied to the data from the ethylene study. The calculations were performed using the SAS® program package.

In case of normality, the maximum likelihood estimates ζ_{ik} coincide with the least squares estimates. Thus, ζ_{ik} can be conveniently estimated using the PROC NLIN procedure. The estimation was performed using the Marquardt algorithm in PROC NLIN (SAS STAT users guide, 1994). The EM algorithm was implemented in SAS/IML®. For programs and further details, see Schirm (1999) and Schirm and Selinski (2000).

The estimates of the kinetic constants are computed in h^{-1}, the initial concentration is measured in *ppm*.

4.1 Estimates for group A

The maximum likelihood or rather least squares estimation of the kinetic parameters and the initial concentration required several weeks. It was possible to obtain estimates for all data sets of group A but for animal 10, 5th occasion, where no observations were available as animal 10 was dropped out of the experiment during the fourth day (cf. Quinke *et al.*, 2000, for further details). The results of the maximum likelihood estimation are given by Schirm and Selinski (2000) and Selinski (2001).

The estimates of the variance τ_{ik}^2, $i = 1, \ldots, 10$, $k = 1, \ldots, 5$, are given by the mean squared residuals. They give a first hint with respect to the fit of the model and the occurrence of possible outliers or some special features of the data, animal 3, 2nd occasion, for instance (see figure 2).

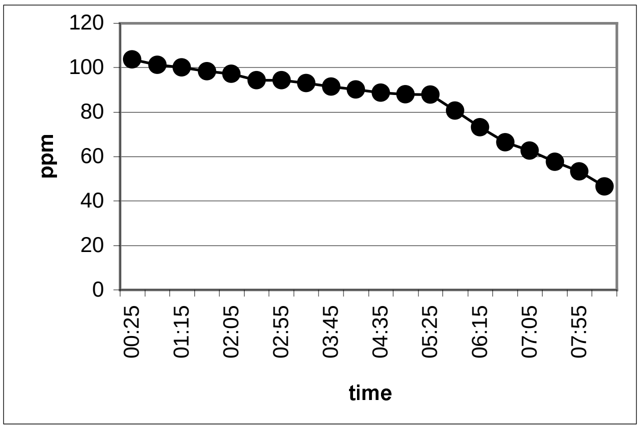

Figure 2. *Measured concentration of ethylene in the atmosphere of the inhalation chamber of rat 3 at the* 2^{nd} *exposure occasion; time in hours since application of ethylene.*

The estimates of τ_{ik}^2 are given in table 1 (Schirm and Selinski, 2000).

Table 1*: Estimates of* τ_{ik}^2 *in group A; occ. denotes the occasion and* J_{ik} *the number of observations.*

rat	occ.	J_{ik}	$\hat{\tau}_{ik}^2$	rat	occ.	J_{ik}	$\hat{\tau}_{ik}^2$
1	1	19	0.3751	6	1	21	1.5588
	2	20	7.6993		2	20	1.8800
	3	21	2.3942		3	21	1.1651
	4	20	1.1378		4	21	1.1731
	5	20	0.8222		5	21	1.0177
2	1	19	0.5984	7	1	20	1.9394
	2	20	0.7304		2	21	0.8443
	3	21	8.0063		3	19	1.5757
	4	21	2.6939		4	21	2.1700
	5	21	8.8819		5	20	1.2412
3	1	19	0.8810	8	1	20	1.4161
	2	20	46.7935		2	19	0.9936
	3	21	0.7284		3	19	0.8557
	4	21	0.7362		4	21	0.4737
	5	19	1.5714		5	20	5.9009

4	1	19	0.3881	9	1	21	0.9357
	2	20	16.7085		2	19	6.3617
	3	21	0.7776		3	19	8.7399
	4	21	0.4518		4	20	2.1776
	5	21	1.0244		5	21	0.3247
5	1	19	0.2953	10	1	21	1.5953
	2	20	0.6920		2	21	0.5121
	3	21	0.5447		3	21	0.7565
	4	20	1.5790		4	16	0.6898
	5	18	0.9940		5	0	---

4.1.1 EM estimation in model A

The EM algorithm as given in section 3.2.2 was implemented using SAS/IML®. The algorithm converged quite fast with computational times of about 10 to 15 minutes. Tables 2 – 4 show the estimators of the population mean, individual mean and specific kinetic constants β, β_i, and β_{ik}, $i = 1, \ldots, 10$, $k = 1, \ldots, 5$ (Schirm and Selinski, 2000).

Table 2. *Estimated population mean parameters from group A.*

k_{12}^*	k_{21}^*	k_{el}^*	$y^*(0)$
0.0195	1.9459	7.9203	120.7751

Table 3. *Estimated individual mean parameters from group A.*

rat	k_{12i}^*	k_{21i}^*	k_{eli}^*	$y_i^*(0)$
1	0.0165	1.7996	8.2271	122.0978
2	0.0251	1.4295	9.7140	125.9349
3	0.0172	1.6769	8.6843	123.6201
4	0.0170	1.5062	9.4373	124.6642
5	0.0395	2.7110	5.9216	109.8376
6	0.0152	1.8898	7.8890	121.3839
7	0.0153	1.7540	8.3365	122.1477
8	0.0185	2.3801	6.5166	116.9047
9	0.0171	2.1202	7.2037	119.2689
10	0.0163	1.9532	7.6405	120.7579

Table 4. *Estimated individual occasion-dependent parameters from group A.*

rat	occasion	k^*_{12i}	k^*_{21i}	k^*_{eli}	$y^*_{ik}(0)$
1	1	0.017867	2.042351	7.813926	120.407321
	2	0.036978	1.177360	8.338775	121.954447
	3	0.027586	3.744748	5.983159	115.581855
	4	0.017331	2.807415	7.212940	118.964772
	5	0.016146	1.754080	8.218957	121.935992
2	1	0.016227	1.813242	8.310704	122.296099
	2	0.015744	0.576604	10.147568	129.166689
	3	0.028768	2.911374	6.958533	118.202267
	4	0.024134	0.983918	8.914000	124.252370
	5	0.027182	1.112956	8.530211	122.715473
3	1	0.013899	1.286805	9.032062	125.334979
	2	0.038556	1.437662	8.196556	121.583943
	3	0.020023	2.689459	6.866247	116.936443
	4	0.020142	2.398732	7.347466	118.715818
	5	0.017125	1.540455	8.297491	122.116976
4	1	0.016811	2.397992	7.108658	118.141629
	2	0.016057	2.093109	7.769620	120.366496
	3	0.018866	2.863103	6.171412	115.147041
	4	0.016776	1.982590	7.929145	120.859000
	5	0.017125	1.253202	9.114037	124.816619
5	1	0.015103	1.372176	8.694388	124.441190
	2	0.016994	2.521933	7.138030	117.399421
	3	0.017026	1.603022	8.333555	122.633271
	4	0.029043	3.919082	5.563675	113.474553
	5	0.049548	3.508707	6.293463	115.809103
6	1	0.015754	1.497588	8.434854	122.763099
	2	0.014801	0.584613	9.353037	126.586386
	3	0.014648	1.596484	8.401046	122.698102
	4	0.015206	1.957088	7.959639	120.984710
	5	0.014473	1.914141	8.028343	121.263291
7	1	0.017682	1.288865	8.488848	122.887635
	2	0.017778	1.817587	8.015805	121.080294
	3	0.015246	1.804568	8.060977	121.285943
	4	0.012670	2.428750	7.394145	118.953311
	5	0.014744	1.695325	8.200570	121.844066

8	1	0.019036	1.201966	8.975108	124.478546
	2	0.018865	1.301862	8.951265	124.449674
	3	0.019677	1.484649	8.703304	123.573169
	4	0.018887	2.643065	6.684446	116.532885
	5	0.017386	2.773638	6.894735	117.743190
9	1	0.018555	2.807156	6.807675	116.768842
	2	0.021778	3.366203	6.798242	118.278075
	3	0.018139	2.209665	7.735031	120.243264
	4	0.016314	1.675937	8.291964	122.270009
	5	0.016378	2.247555	7.544129	119.360711
10	1	0.021022	1.450429	8.325077	122.279515
	2	0.026133	2.643840	6.920083	116.773079
	3	0.020052	3.172268	6.344221	115.066653
	4	0.015975	2.016850	7.893261	120.738310

These results are consistent with the maximum-likelihood estimates. In general, extreme data points in some components are corrected towards a common mean by the Bayes estimation.

A comparison of the interindividual and interoccasional variability can be made by computing estimates of the covariance matrices Ω_i, $i = 1, \ldots, 10$ and Σ.

Table 5. *Estimates of the individual covariance matrices* Ω_i *from group A.*

rat	k_{12}	k_{21}	k_{el}	$y(0)$
1	1.1112	0.0011	-0.0027	-0.0088
	0.0011	1.6942	-0.6174	-1.7940
	-0.0027	-0.6174	1.8052	2.0536
	-0.0088	-1.7940	2.0536	7.2421
2	1.1111	0.0011	-0.0004	-0.0035
	0.0011	1.4855	-0.4734	-1.5382
	-0.0004	-0.4734	2.4212	3.6635
	-0.0035	-1.5382	3.6635	11.8527
3	1.1112	0.0001	-0.0023	-0.0092
	0.0001	1.3083	-0.3080	-1.1427
	-0.0023	-0.3080	1.7335	2.3199
	-0.0092	-1.1427	2.3199	9.7854

4	1.1111	0.0002	-0.0005	-0.0014
	0.0002	1.4746	-0.9026	-2.5671
	-0.0005	-0.9026	3.4722	6.5698
	-0.0014	-2.5671	6.5698	19.5659
5	1.1113	0.0064	-0.0157	-0.0880
	0.0064	1.6835	-0.7500	-2.8888
	-0.0157	-0.7500	2.8058	9.0527
	-0.0880	-2.8888	9.0527	54.7850
6	1.1111	0.0000	0.0000	-0.0002
	0.0000	1.3276	-0.2519	-0.8607
	0.0000	-0.2519	1.4142	0.9997
	-0.0002	-0.8607	0.9997	4.5410
7	1.1111	-0.0003	0.0002	0.0008
	-0.0003	1.1869	-0.0815	-0.2881
	0.0002	-0.0815	1.2343	0.4160
	0.0008	-0.2881	0.4160	2.5251
8	1.1111	-0.0003	0.0005	0.0014
	-0.0003	1.5085	-0.8097	-2.5331
	0.0005	-0.8097	2.9916	5.7585
	0.0014	-2.5331	5.7585	18.8444
9	1.1111	0.0008	-0.0003	-0.0011
	0.0008	1.3607	-0.1300	-0.4652
	-0.0003	-0.1300	1.3226	0.5785
	-0.0011	-0.4652	0.5785	3.0218
10	1.1111	0.0010	-0.0010	-0.0059
	0.0010	1.3578	-0.2673	-1.1618
	-0.0010	-0.2673	1.4147	1.2539
	-0.0059	-1.1618	1.2539	6.7316

Table 6. *Estimates of the population covariance matrix* Σ *from group A.*

	k_{12}	k_{21}	k_{el}	$y(0)$
k_{12}	1.1112	0.0016	-0.0038	-0.0235
k_{21}	0.0016	1.2665	-0.4653	-1.7714
k_{el}	-0.0038	-0.4653	2.5475	5.1850
$y(0)$	-0.0235	-1.7714	5.1850	22.3028

Interestingly, all covariance matrices Ω_i, $i = 1, \ldots, 10$, are all very similar and they are very similar to the covariance matrix Σ. They have the same structure and all very similar entries. Only the variance of the initial concentration differs much between the animals. The rates of uptake, exhalation, and metabolism seem to be independent. Furthermore, there is no clear dependency between the initial concentration and the rates of uptake and exhalation. On the contrary the results indicate a correlation between the rate of elimination and initial concentration of ethylene, which is not consistent with the assumption of a first order kinetic process.

To evaluate the quality of the estimates the coefficient of determining

$$R^2 = 1 - \sum_{j=1}^{J} \left(y_j - \tilde{y}_j \right)^2 \Big/ \sum_{j}^{J} \left(y_j - \bar{y}_j \right)^2 \qquad (27)$$

is calculated, where y_j denote the observations and the $\tilde{y}_j$ are the estimated observations. R^2 provides a measure of fit of the model relating the variance explained by the model to the total variance.

R^2 was calculated for all single inhalation experiments separately and for all estimates including the maximum likelihood estimates. For the individual mean and the population mean parameters the initial concentration from β^* was used. Otherwise the estimated individual and population mean kinetic parameters would appear to be rather bad just due to the shift of the estimated concentration-time curves to a higher or lower initial concentration than the one of the specific experiment. Note, that the initial concentration is only a scaling factor in case of kinetic processes of first order (Selinski, 2001).

***Table** 7. R^2 calculated for the maximum likelihood estimates ζ_{ik} and for the Bayes estimates β^*, β^*, β^* from group A as well as the difference between $R^2(\beta^*)$ and $R^2(\zeta_{ik})$.*

rat	occasion	$R^2(\zeta_{ik})$	$R^2(\beta^*)$	$R^2(\beta^*)$	$R^2(\beta^*)$	difference
1	1	0.9933	0.9933	0.8813	0.9784	0.0000
	2	0.9693	0.9685	0.1121	0.4478	-0.0008
	3	0.9667	0.8113	-5.1731	-3.4714	-0.1554
	4	0.9736	0.8248	-0.6271	0.3758	-0.1488
	5	0.9873	0.9869	0.9801	0.7243	-0.0004
2	1	0.9877	0.9873	0.3461	0.7131	-0.0005
	2	0.9900	0.9782	-1.2830	-0.8462	-0.0118
	3	0.9201	0.8446	-0.6997	-0.9633	-0.0755
	4	0.9806	0.9437	0.8467	0.8360	-0.0369
	5	0.9479	0.8985	0.9188	0.8456	-0.0494

3	1	0.9796	0.9698	0.3083	-0.9110	-0.0099
	2	0.8358	0.8325	0.0078	0.3242	-0.0032
	3	0.9873	0.9771	-0.3935	0.4591	-0.0102
	4	0.9886	0.9858	0.2485	0.8083	-0.0028
	5	0.9746	0.9440	0.9259	0.5036	-0.0306
4	1	0.9910	0.9898	0.1595	0.9097	-0.0012
	2	0.7453	0.7450	0.5755	0.6727	-0.0003
	3	0.9838	0.9701	-1.6673	0.0394	-0.0137
	4	0.9924	0.9924	0.8422	0.9078	0.0000
	5	0.9870	0.9829	0.9649	0.4753	-0.0041
5	1	0.9945	0.9944	-11.0791	0.1413	-0.0002
	2	0.9857	0.9843	-4.8625	0.8815	-0.0014
	3	0.9925	0.9925	-6.2837	0.7823	0.0000
	4	0.9751	0.7795	0.6418	-5.2533	-0.1956
	5	0.9921	0.6241	-0.4460	-7.1038	-0.3680
6	1	0.9759	0.9695	0.8885	0.4011	-0.0065
	2	0.9746	0.9298	-0.1257	-1.2145	-0.0447
	3	0.9786	0.9759	0.8493	0.1946	-0.0027
	4	0.9773	0.9773	0.9767	0.6932	0.0000
	5	0.9788	0.9786	0.9659	0.5065	-0.0003
7	1	0.9775	0.9668	0.9441	0.6265	-0.0106
	2	0.9887	0.9886	0.8872	0.9291	0.0000
	3	0.9755	0.9753	0.9743	0.6397	-0.0002
	4	0.9397	0.9281	0.7094	0.5337	-0.0117
	5	0.9781	0.9772	0.9633	0.3780	-0.0010
8	1	0.9858	0.9792	0.3057	0.7277	-0.0066
	2	0.9898	0.9853	0.3516	0.7688	-0.0045
	3	0.9914	0.9892	0.6031	0.9049	-0.0022
	4	0.9910	0.9873	0.9435	0.5890	-0.0037
	5	0.8826	0.8439	0.8107	0.3617	-0.0388
9	1	0.9824	0.9587	0.3894	0.4211	-0.0237
	2	0.8852	0.6014	-2.1123	-1.9223	-0.2839
	3	0.8843	0.8805	0.8685	0.8673	-0.0038
	4	0.9647	0.9600	0.6595	0.6196	-0.0046
	5	0.9937	0.9932	0.9907	0.9452	-0.0005
10	1	0.9854	0.9828	0.9220	0.9523	-0.0027
	2	0.9949	0.9823	-0.4063	0.1020	-0.0126
	3	0.9862	0.9516	-1.1030	-0.3694	-0.0346
	4	0.9811	0.9811	0.9796	0.8360	0.0000
median		0.9811	0.9759	0.6031	0.5337	-0.0041
minimum		0.7453	0.6014	-11.0791	-7.1038	-0.3680
maximum		0.9949	0.9944	0.9907	0.9784	0.0000

In general the fit of the individual and occasion specific parameter vectors β^* is very good. The median of R^2 is about 0.98, it minimum value is about 0.60 and its maximum about 0.99. In five cases the performance of the Bayes estimate is much worse than the fit of the maximum likelihood estimate. The difference in the worst case is 0.3680 (rat 5, 5th occasion). Considering the fit of the individual mean and the population mean yield differing results. R^2 ranges from 0.99 to –11.08 in case of the individual mean and from 0.98 to –7.10 in case of the population mean. In most cases, the fit of both means to the data from the single experiments is satisfying. A discussion and an analysis of the outlier problem is given by Selinski (2001) and Selinski and Becker (2001).

4.2 Estimates for group B

Analysing the data from group B (repeated application of different doses) yielded poor maximum likelihood and Bayes estimates. Due to the bad performance of these estimates, outliers were detected and eliminated from the data. The maximum likelihood and EM estimation were repeated afterwards. As no animal was dropped out of the experiment the data set is complete.

The maximum likelihood or rather least squares estimation of the kinetic parameters and the initial concentration was performed using the Marquardt algorithm in PROC NLIN (SAS STAT users guide, 1994; Schirm, 1999). The estimation procedure required even more time than for group A. Although several sets of starting values were used it was not possible to obtain estimates for all individuals and dosing occasions. The algorithm did not converge for rat 11, 2nd and 5th dose and for rat 16, 5th dose (Selinski, 2001).

The EM algorithm was implemented using SAS/IML® (Schirm, 1999; Schirm and Selinski, 2000) and required 15 iterations. So, the computational effort was minor compared with the least squares estimation. Note, that there are no estimates of φ_{ik} available for rat 11, 2nd and 5th dose, and for rat 16, 5th dose as the respective maximum likelihood estimates were not available.

Although the performance of the maximum likelihood estimates was good – median = 0.97 – the fit of the individual and dose specific parameter vectors $\boldsymbol{\varphi}_k^*$, $i = 11, \ldots, 20$, $k = 1, \ldots, 5$, is very poor for several animals (cf. Selinski, 2001 for further details). Rat 12, 13, 15, 17, and 20, where the rates of uptake were all negative, showed values of R^2 from -3585 to -47 Millard! For the rest of the animals the results were very good. The lack of fit was not related to the body weight of the individuals. Considering the individual and population mean the performance of the population mean is much better than of the individual mean and – in case of the rats with poor fit of φ_k^* – much better than the performance of the individual and dose-dependent parameter vectors. Note, that there is no relationship between $R^2(\varphi_i)$ and $R^2(\varphi_{ik})$, i.e. animals with a poor fit of φ_k^* do not display worse results for φ^* than the other individuals.

The following observations were identified as outliers using a modified Hampel identifier (Selinski and Becker, 2001).

Table 8. *Outliers in group B, estimation of the concentration-time curve performed by the use of the Bayes estimate of ζ_{ik} from group B, time in hours since application of ethylene.*

rat	occasion	time	rat	occasion	time
11	3	0:25	15	1	3:20
11	3	7:05	17	3	3:20
12	1	0:25	18	1	4:10
13	2	8:20	18	3	2:30
13	5	0:25	18	3	7:30
14	1	3:45	18	5	3:20
14	2	0:25	20	1	5:00
14	3	3:20	20	5	0:25
14	3	3:45	20	5	5:25
14	4	375	20	5	500

After removing the observations given by table 8 from the data set and checking the assumptions of first order kinetics the estimation procedure was repeated. Using several sets of starting values the least squares estimation took several weeks. Again, it was not possible to obtain estimates for rat 11 and 16, 5th dosing occasion, both (Selinski, 2001).

As no maximum likelihood estimates for rat 11 and 16, 5th dose, were available the population parameters from the first EM estimation were used instead. The estimates of the kinetic parameters were substituted by .00414 for the rate of uptake, 1.27489 for the exhalation, and 8.35140 for the metabolic elimination. For details with respect to the estimation of the initial concentrations of both experiments see Selinski (2001). Thus, τ_k^2 was estimated for these particular data sets as $\hat{\tau}_{11,5}^2$ = 21.1392

($J_{11,5}$ = 20) and $\hat{\tau}^2_{16,5}$ = 10.9553 ($J_{16,5}$ = 21). The estimates of τ_k^2 for the rest of the data sets are given in table 9.

Table 9. *Estimates of the variance* τ_k^2 *and numbers of observations* J_{ik} *(second estimation).*

rat	dose	J_{ik}	$\hat{\tau}_k^2$	rat	dose	J_{ik}	$\hat{\tau}_k^2$
11	1	21	0.0664	16	1	21	0.1465
	2	21	0.5956		2	21	0.4888
	3	19	0.5426		3	21	1.4759
	4	21	2.5780		4	21	6.1062
	5	–	–		5	–	–
12	1	20	0.0142	17	1	21	0.1074
	2	21	0.6108		2	21	0.2679
	3	21	1.0455		3	20	0.6135
	4	20	2.3203		4	21	3.2113
	5	20	32.9998		5	21	27.5246
13	1	21	0.1999	18	1	19	0.0839
	2	19	0.2281		2	19	0.1962
	3	21	0.2952		3	18	0.2688
	4	20	1.4387		4	21	3.8432
	5	20	11.7717		5	20	11.1034
14	1	20	0.0420	19	1	20	0.0871
	2	20	0.3239		2	21	0.4350
	3	18	0.2687		3	21	1.6764
	4	19	2.4961		4	21	9.2073
	5	21	22.1548		5	21	14.4091
15	1	20	0.0885	20	1	20	0.1011
	2	21	0.3462		2	21	0.1985
	3	21	0.8770		3	21	1.7925
	4	21	1.6208		4	21	5.4733
	5	21	20.9219		5	18	3.2550

A second EM estimation was performed with the new maximum likelihood estimates. The EM algorithm converged within 8 minutes requiring 21 iterations. The new population mean is given in table 10. Note, that the estimated rates of uptake and exhalation are higher than those from the first estimation procedure, whereas the rate of metabolism is lower.

Table 10. *Bayes estimates of the population mean* φ *of the kinetic parameters from group B (second estimation).*

k_{12}^*	k_{21}^*	k^*
0.00919383	2.48195581	6.07420137

The second estimates of the individual means contain no negative rates of uptake. The new estimated rates of exhalation are higher than the respective estimates from the first estimation procedure whereas the new estimates of the rates of metabolism are lower (see table 11).

Table 11. *Bayes estimates of the individual means* φ_i *of the kinetic parameters from group B (second estimation).*

rat	k_{12i}^*	k_{21i}^*	k_i^*
11	0.008159	2.566444	7.955422
12	0.004663	2.632562	4.996317
13	0.009752	2.703144	4.963009
14	0.011046	2.619991	3.841387
15	0.007859	1.522446	4.381336
16	0.010371	2.764825	4.164892
17	0.006331	3.091274	6.510701
18	0.010644	2.397212	3.208884
19	0.007144	2.589783	11.35654
20	0.008846	1.839113	7.420888

The estimates of the individual and dose specific kinetic parameters are given in table 12. For 4 animals negative estimates of the rate of uptake were obtained. These were rat 12, all but 3rd dose, rat 13, only 3rd dose, rat 14, 1st and 4th dose, and rat 16, all doses. Remarkably, these individuals correspond only partly to those of the first estimation procedure whose estimates of k_{12ik} were also negative (rats 12, 13, 15, 17, and 20). Furthermore, not all of the estimated rates of uptake of these specific individual were affected.

Table 12. *Bayes estimates of the individual and dose-specific kinetic parameters* φ_{ik} *from group B (second estimation).*

rat	dose	k^{*}_{12ik}	k^{*}_{21ik}	k^{*}_{ik}
11	1	0.00798	2.585867	6.206487
	2	0.008014	2.561593	6.202799
	3	0.008203	2.263459	6.275936
	4	0.007949	2.59943	6.20916
	5	0.007954	2.389376	6.217159
12	1	-0.00367	2.721866	5.968923
	2	-0.00284	2.821971	5.935779
	3	0.002171	2.701235	5.971985
	4	-0.00025	2.845661	5.923158
	5	0.002122	2.761023	5.95977
13	1	0.009891	2.824627	5.738727
	2	0.004327	2.814266	5.741295
	3	-0.00105	3.060259	5.547908
	4	0.004996	2.858488	5.727517
	5	0.005535	2.931295	5.688083
14	1	-0.01011	2.885706	5.606174
	2	0.004299	2.739558	5.662993
	3	0.003109	2.743573	5.662277
	4	-0.00286	3.27717	5.038739
	5	0.001784	2.842023	5.630116
15	1	0.08689	0.926819	4.520482
	2	0.033398	1.88046	5.057021
	3	0.033398	1.88046	5.057021
	4	0.033398	1.88046	5.057021
	5	0.03354	0.900386	3.704453
16	1	-0.01726	3.050767	5.75792
	2	-0.0008	2.898646	5.830125
	3	-0.00297	3.027369	5.773353
	4	-0.0031	3.251369	5.565021
	5	-0.00197	3.125021	5.701347
17	1	0.004988	3.396908	6.391919
	2	0.004982	3.398874	6.394717
	3	0.004981	3.400052	6.596803
	4	0.005026	3.361487	6.353749
	5	0.004986	3.155935	6.317406
18	1	0.009444	3.057763	5.02804
	2	0.011858	2.338577	5.688958

	3	0.011769	2.414216	5.705137
	4	0.010302	2.530091	5.697991
	5	0.013028	2.306737	5.677196
19	1	0.006098	2.700983	6.733596
	2	0.006085	2.706705	6.738767
	3	0.006562	2.327167	6.597734
	4	0.006732	1.757249	7.211028
	5	0.006167	1.401046	8.387891
20	1	0.008978	1.490754	6.164023
	2	0.008966	1.4981	6.159701
	3	0.00897	1.491188	6.163765
	4	0.008975	1.462438	6.181617
	5	0.008964	1.498669	6.15937

The estimates of the covariance matrices in model B provide information about the intraindividual differences under different exposure conditions (see table 13) and the interindividual variability of the population of test animals (see table 14).

Table 13. *Estimates of the intraindividual covariance matrices* $\Omega_{11}, \ldots, \Omega_{20}$ *from group B (second estimation).*

rat	k_{12}	k_{21}	k_{el}
11	2.500000	0.000003	0.000304
	0.000003	2.531161	0.183396
	0.000304	0.183396	6.255533
12	2.500014	-0.000652	0.005148
	-0.000652	2.550923	-0.237880
	0.005148	-0.237880	4.379745
13	2.500004	0.000471	-0.002723
	0.000471	2.584080	-0.355338
	-0.002723	-0.355338	4.483218
14	2.500011	0.000518	-0.009002
	0.000518	2.546324	-0.420793
	-0.009002	-0.420793	9.586907

15	2.500000	0.000181	0.000374
	0.000181	3.667813	2.199416
	0.000374	2.199416	6.737400
16	2.500007	0.000825	-0.006043
	0.000825	2.623253	-0.736641
	-0.006043	-0.736641	7.792127
17	2.500004	-0.001303	-0.000606
	-0.001303	2.988963	0.216383
	-0.000606	0.216383	2.604890
18	2.500009	-0.000283	-0.009880
	-0.000283	2.530364	0.307641
	-0.009880	0.307641	13.851830
19	2.500001	-0.000132	-0.005617
	-0.000132	2.536888	0.700922
	-0.005617	0.700922	35.451378
20	2.500001	-0.000673	0.001241
	-0.000673	3.035140	-0.964107
	0.001241	-0.964107	4.296670

Table 14. *Estimates of the interindividual covariance matrix Σ from group B (second estimation).*

	k_{12}	k_{21}	k_{el}
k_{12}	1.111116	-0.000071	-0.002056
k_{21}	-0.000071	1.322422	0.082168
k_{el}	-0.002056	0.082168	7.283284

As in case of group A the intra- and interindividual covariance matrices are all very similar. They have the same structure and very similar entries. Regarding just the intraindividual covariance matrices $\Omega_{11}, \ldots, \Omega_{20}$ shows that the differences in the variability structure of the processes between the individuals manifest most in the metabolic elimination. The variance of k_{elik} and the covariance of k_{elik} and k_{21ik}, the rate of exhalation show the most dissimilarity between the individuals. In general, the kinetic constants seem to be independent as required in case of overall first order kinetics.

According to the proceeding for the first estimation the fit to the data is measured in terms of R^2. Database for the calculation of R^2 is the data set without the outliers given in table 8 The results are given in table 15. Note, that no maximum likelihood estimates were available for rat 11 and 16, 5th dose.

Table 15. *R^2 calculated for ζ_{ik}, φ_{ik}, φ_i, and φ, data set without outliers.*

rat	dose	$R^2(\zeta_{ik})$	$R^2(\varphi_{ik})$	$R^2(\varphi_i)$	$R^2(\varphi)$
11	1	0.9070	0.1364	0.8925	-0.1906
	2	0.9659	0.5335	0.0426	0.6780
	3	0.9835	-1.4506	-2.1779	-0.3693
	4	0.9747	-6.6341	-3.7872	-7.2245
	5	–	-30.6329	-21.4582	-40.3800
12	1	0.5698	-56.0277	-0.3866	-5.2208
	2	0.9745	-7.9342	-3.0105	-2.9447
	3	0.9833	0.0187	-0.4433	-0.0244
	4	0.9806	-2.6260	-0.1046	0.0405
	5	0.8254	-19.9743	-33.0128	-30.1312
13	1	0.7619	0.4046	0.6832	-0.8849
	2	0.9825	0.3017	-0.6175	0.5237
	3	0.9920	-14.8619	-5.4395	-9.9383
	4	0.9860	-0.2184	-2.1770	0.0048
	5	0.8978	-51.8458	-80.9596	-49.3104
14	1	0.9427	-47.2022	-0.7459	-11.8703
	2	0.9792	0.2899	-0.4865	0.7275
	3	0.9964	0.2759	-0.6973	0.6463
	4	0.9945	-4.3776	-0.9957	-3.2818
	5	0.9315	-4.1971	-24.8273	-6.8534
15	1	0.8400	-165.3988	-0.4292	-2.8251
	2	0.9754	-7.9177	0.5076	0.0214
	3	0.9839	-5.6725	0.7207	0.5690
	4	0.9827	-8.3406	0.4230	0.9108
	5	0.9153	-97.0310	-22.4543	-32.1130
16	1	0.8251	-70.2370	-0.5869	-10.4834
	2	0.9686	-0.3063	-3.3844	-0.3461
	3	0.9785	-0.5551	-1.1147	0.6872
	4	0.9359	-1.4125	-9.1157	-1.5358
	5		-18.8334	-83.1219	-32.1821

17	1	0.5318	0.4865	0.4495	-0.0382
	2	0.9775	-1.5767	-1.1191	-0.8332
	3	0.9903	-2.3528	-2.4726	-2.4384
	4	0.9690	-4.3954	-5.0761	-5.1842
	5	0.8448	-9.6885	-9.6459	-8.7681
18	1	0.8887	-18.7871	-9.0275	-38.0404
	2	0.9778	-1.1567	-6.2555	-0.3893
	3	0.9958	0.8797	-0.3047	0.6762
	4	0.9651	0.8667	-0.7362	0.5556
	5	0.9255	-67.0190	-117.8679	-50.5485
19	1	0.7689	-5.3373	-0.2019	-9.6235
	2	0.9678	-0.2597	0.6302	-0.7742
	3	0.9904	-0.7991	-1.5091	-0.1413
	4	0.9671	-2.3346	-2.3890	-0.0730
	5	0.9763	0.0518	0.8224	-0.5576
20	1	0.8686	0.7122	0.6863	-2.1347
	2	0.9833	0.6689	0.6524	-0.5648
	3	0.9772	0.6573	0.6519	0.2906
	4	0.9666	0.7684	0.7712	0.1563
	5	0.9837	-15.3233	-15.5692	-34.6259
median		0.9688	-2.3437	-0.8708	-0.8037
minimum		0.5318	-165.3988	-117.8679	-50.5485
maximum		0.9964	0.8797	0.8925	0.9108

In general, the performance of the maximum likelihood estimates is quite good (median = 0.9688) whereas the fit of φ_{ik}^{*} is very poor for most individuals and doses. R^2 is ranging from -165 to 0.8797 for φ_{ik}^{*}. The performance of the individual and the population mean is similar though their fit is in general better than of φ_{ik}^{*} (median = -0.87 and -0.80, respectively). Remarkably, the fit of the different Bayes estimates to the single data sets is often discordant, i.e., poor fit of φ_{ik}^{*}, worse fit of φ^{*} and good fit of φ_{i}^{*}. For a detailed discussion of the results and the estimation procedure, see Selinski, 2001.

4.3 Checking assumptions of first order kinetics

The experiments of group B, where ten rats were exposed five times each to different concentrations of ethylene of about 20, 50, 100, 200, and 500 *ppm* ethylene, where used to check the assumption of ’good’ approximation of the real kinetic processes by first order kinetics applying Kendall’s test of independence .

To test the null hypothesis of independence of the standardised observations from the initial concentration the 18th observation, after 7 ½ hours, had been chosen as one of the last time points where the observations are complete. The maximum likelihood estimates of the initial concentrations (first estimation procedure) were used to standardise the observations (Selinski, 2001). Alternatively the first observations may be used. Thus, it is possible to test the hypothesis of overall first order kinetics without fitting a model.

Furthermore, checking the single processes the respective estimates of the standardised individual and occasion dependent rates of uptake, exhalation, and metabolism and the corresponding estimated initial concentrations where used.

Moreover, Kendall's modified correlation coefficients τ^B are calculated.

4.3.1 Test of overall first order kinetic processes

Using the maximum likelihood estimates of the initial concentrations and the large-sample approximation in the case of ties (26) the Kendall correlation coefficient was given by $\tau^B = 0.14122$ for the 18th observations.

The probability of the given value of the test statistic under the null hypothesis, the so-called *p-value*, is 0.1479.

Estimating the initial concentration by the first observation yielded $\tau^B = 0.10367$ corresponding to a p-value of 0.2881.

Thus, the null hypothesis of independence of the standardised observations from the initial concentration could not be rejected for neither choice of the estimate for the dose. Hence, overall first order kinetic processes seem to be a good approximate of the real kinetic processes of uptake, exhalation, and metabolism of ethylene.

4.3.2Test of first order kinetic partial processes

The results of the corresponding tests and Kendall correlation coefficients for each partial process are given in table 16:

Table 16. *Results of Kendall's test for independence and Kendall's correlation coefficients.*

partial process	τ^B	p-value
uptake (k_{12ik})	0.05735	0.5581
exhalation (k_{21ik})	0.01143	0.9068
metabolism ($k_{el\,ik}$)	0.08657	0.3752

Thus, the null hypothesis of independence of the respective individual and occasion dependent rate and the initial concentration cannot be rejected. Hence, considering all partial processes separately, the first order approximation of the kinetic processes seems to be valid.

4.3.3 Graphical analysis

The estimation of $y_2(t_{\max})$ is performed using φ_{ik} and the maximum likelihood estimates $\hat{y}_{ik}(0)$ for the comparison with the Bayes estimates of k_{21ik} and k_{elik} and using ζ_{ik} for the comparison with the respective maximum likelihood estimates. Thus, we compared all estimates, which may possibly reveal a dependency between parameters and concentrations.

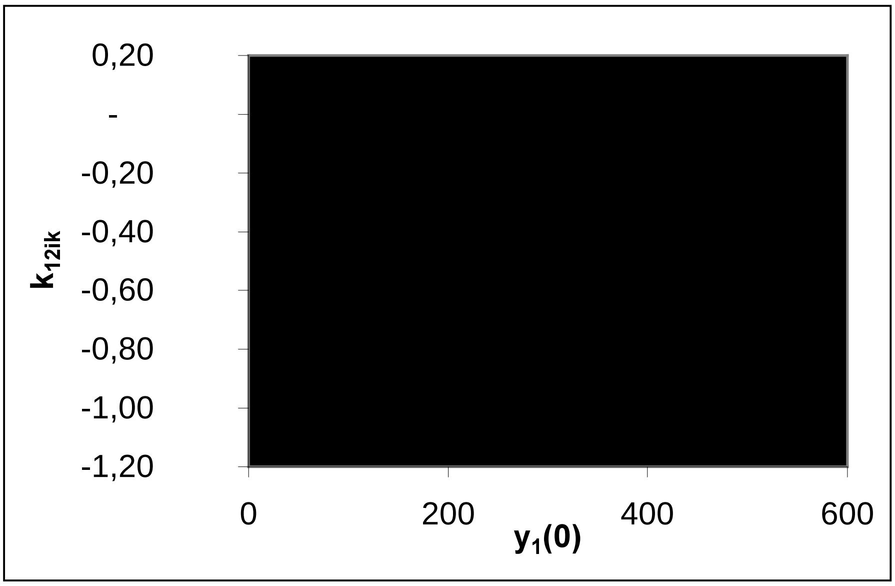

Figure 3. *Comparison of Bayes estimates of k_{12ik} and estimated maximum concentration in the first compartment.*

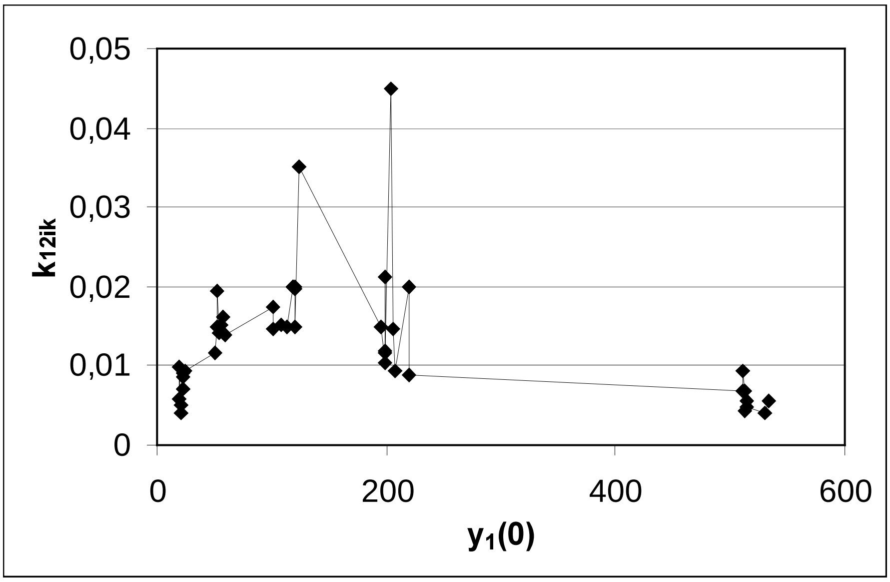

Figure 4. *Comparison of maximum likelihood estimates of k_{12ik} and estimated maximum concentration in the first compartment.*

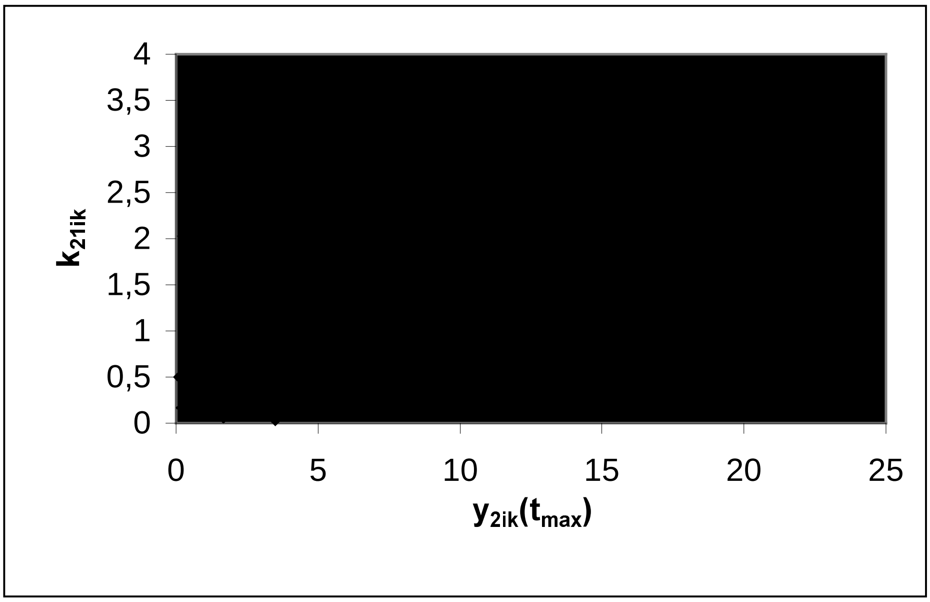

Figure 5. *Comparison of Bayes estimates of k_{21ik} and estimated maximum concentrations in the second compartment using φ_{ik} and the maximum likelihood estimates of $y_{ik}(0)$.*

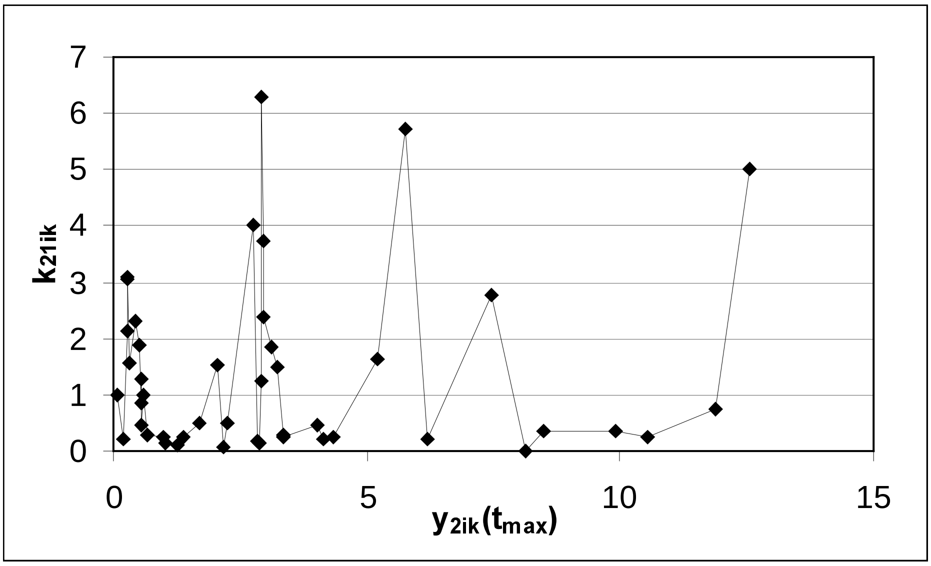

Figure 6. *Comparison of maximum likelihood estimates of k_{21ik} and estimated maximum concentrations in the second compartment using ζ_{ik}.*

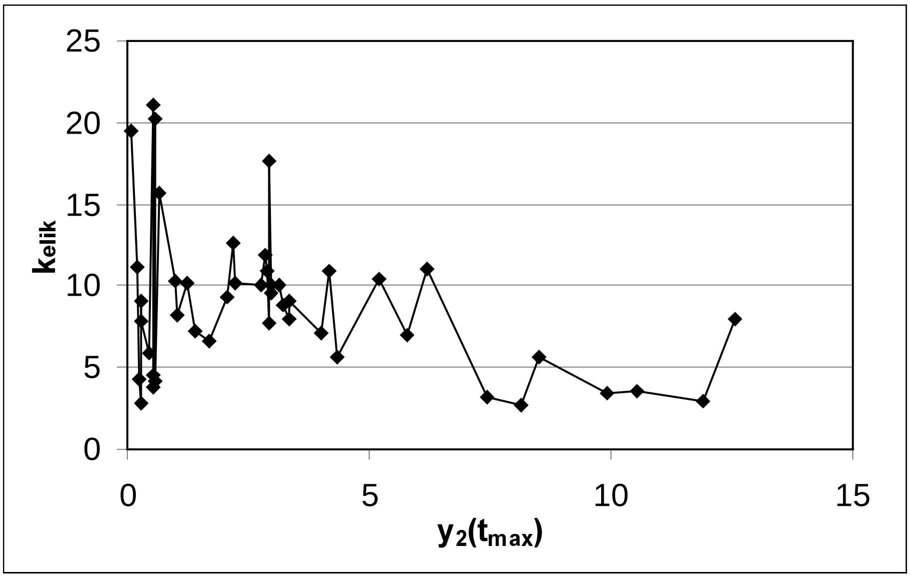

***Figure** 7. Comparison of Bayes estimates of k_{elik} and estimated maximum concentrations in the second compartment using φ_{ik} and the maximum likelihood estimates of $y_{ik}(0)$.*

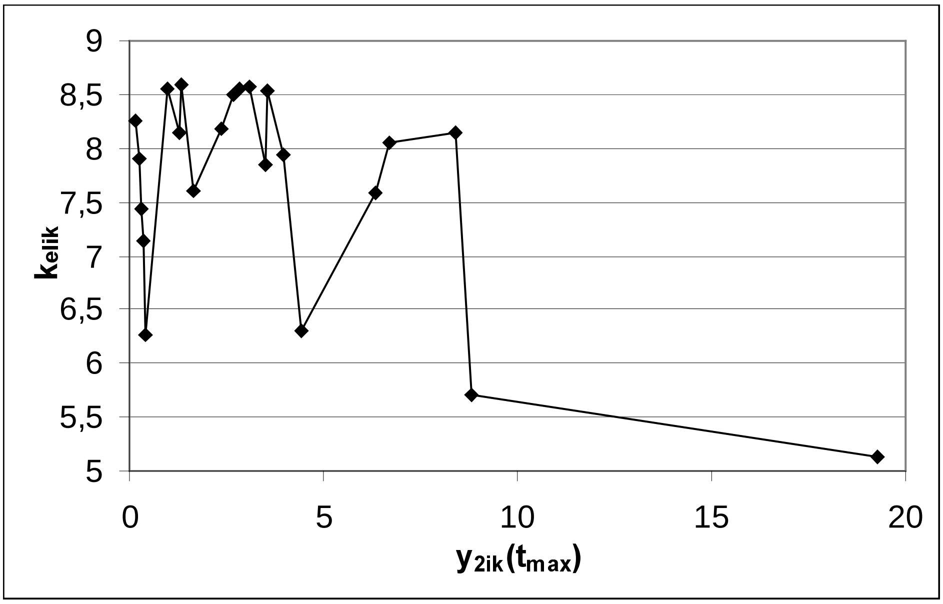

***Figure 8.** Comparison of maximum likelihood estimates of k_{elik} and estimated maximum concentrations in the second compartment using ζ_{ik}.*

The graphical analysis (see figure 3 – 8) supports the validity of the assumption of overall first order kinetics. No graph revealed a dependency or a clear structure so that a functional relationship between the estimated kinetic parameters and the maximum concentration in the respective compartments could not be detected.

Thus, we conclude that the real kinetic processes of uptake, exhalation, and metabolic elimination of ethylene may be well approximated by first order kinetics up to concentrations of about 500 *ppm* in inhalation studies of the described design.

4.4 Conclusions

The disadvantage of the presented method is the duration of the least squares estimation. The subsequent application of the EM algorithm, which converged in general within a few minutes, requires therefore a computational effort, which could be neglected. The Bayes estimation via the EM algorithm provides important information about the covariance structure of the processes and, at least in case of model A, valid population parameters of the investigated processes.

The performance of the Bayes estimates in model A is satisfying so that a reanalysis after elimination of outliers as performed for data set B is omitted. Thus model A yields estimates of the kinetic parameters that seem to be valid for the whole population of test animals.

The comparison of the results of the first and second estimation of model B is quite difficult. For several animals, where the performance of the estimated individual dose specific parameters was very good, the second EM estimation yielded worse results than the first estimation procedure. On the other hand the estimates for the rest of the individuals were not as severe as in case of the first estimation procedure. Nevertheless, the performance of the Bayes estimates in model B is very poor. It can be excluded that the lack of fit of the first estimates of φ_{ik} for several animals is caused by endogenous factors as the respective second estimates were not remarkable. Probably the lack of fit of φ_i^* is due to numerical difficulties related to the inversion of the respective matrices. Although the outlier identifying procedure worked well (Selinski and Becker, 2001), a positive effect of the elimination of outliers and a subsequent reanalysis cannot be deduced from the present example and has to be investigated using different models and / or data sets.

The main results of the present analysis are the population parameters for the processes of uptake, exhalation, and metabolism of ethylene for male Sprague-Dawley rats and the estimation of individual and population covariance matrices. The latter provide useful information about the variability of the investigated processes within the population under investigation.

The estimates of the population mean from the different models and estimation procedures vary within the range of the individual outcomes of the experiments. For the rate of uptake the estimates of group B are lower than for group A. The estimates of group A are higher than the value given in the literature (0.0111 h^{-1}; Filser and Bolt, 1984), the estimate from model B (second estimation) is almost the same. In case of the exhalation and metabolism the results cannot be ordered according to the design of the study. Except the second estimate from model B they are also higher than the values given in the literature for the exhalation and metabolism of ethylene (0.37 h^{-1} and 6.95 h^{-1}, respectively) (Filser and Bolt, 1984).

Considering the intraindividual covariance matrices $\Omega_1, \dots, \Omega_{20}$ from both experimental designs reveals that the rate of metabolism displays more differences between the individuals than the rates of uptake and exhalation do. The latter describe the interaction of the organism with its environment whereas the metabolic elimination of the substance within the organism is expected to be less influenced by environmental factors. Thus, differences between the individuals should manifest in that particular process.

Though tests of independence and graphical analysis suggest that the approximation of the real kinetic processes by first-order kinetics is valid, the estimated covariance matrices of group A give a hint to a possible violation of the assumptions of overall first order kinetics as the rate of metabolism and the initial concentration show a slight dependency.

The similarity of the individual covariance matrices with each other and with the population covariance matrix, which is apparent for both data sets is probably due to the close genetic relationship of the animals as inbreed strains are used for experimentation.

5. Discussion

The main results of the present evaluation are the similarity of the interoccasion covariance matrices with one another and their similarity to the intersubject covariance matrix. It means that the individual mean processes are varying across the population mean process much in the same way as the occasion-dependent processes of all individuals do across their respective individual mean processes. This is obviously related to the close genetic relationship of inbred test animals.

Furthermore, the individual differences seem to manifest more in the metabolic elimination than in the uptake and exhalation, processes that are influenced more by environmental factors than the metabolism.

The present approach simplifies the complex biological processes of highly organised living organisms by the reduction to two-compartment models and the approximation of non-linear kinetics by linear ones. Using linear kinetics we have to be aware of the possible errors, which result from the dependence of the parameters on the concentration if the underlying processes are non-linear. Assuming first order kinetics the processes of uptake, exhalation, and metabolic elimination are independent from the dose. Before summarising the information provided by experiments within a range of concentrations, like in the experiments of group B, it is necessary to verify that a first order approximation of the processes is valid. In fact, the experiments of group A show a correlation between the metabolism and the initial concentration. However, a careful analysis of the data of group B, where a possible violation of the assumptions of first order kinetics should become apparent, confirmed that the approximation of the real kinetic processes by first order kinetics seems to be valid up to concentrations of about 500 *ppm* ethylene.

Implementing the models in a computer using SAS/IML®, we experienced severe numerical difficulties, especially with model B and more complex approaches (Selinski, 2001).

Model A, while neglecting some aspects of the covariance structure of the parameter vectors, has the advantage to be computable by a numerically stable algorithm and therefore yielding numerically quite accurate results.

Due to the lack of fit of the first and second estimates of model B using an alternative approach, Markov Chain Monte Carlo methods, for instance, may provide a methodology, which enable the estimation of all relevant parameters as well as the estimation of the intra- and interindividual covariance matrices. Thus it seems likely to avoid the computational difficulties which occurred with the present approach.

Recent research on Gibbs sampling, a Markov Chain Monte Carlo method closely related to the EM algorithm, has great potential for estimating the parameters of complex models, because it reduces the problem of dealing simultaneously with a large number of related parameters into a much simpler problem of dealing with one unknown quantity at a time. A further application of the EM algorithm in genetics is given by Urfer *et al.* (1999). Gilks *et al.* (1993) have reviewed applications of Gibbs sampling in immunology, pharmacology, cancer screening, industrial and genetic epidemiology. Wikle *et al.* (1998) propose the use of hierarchical Bayesian space-time model with five stages to achieve more flexible models and methods for the analysis of environmental data distributed in space and time. They implement their models in a Markov chain Monte Carlo framework using the Gibbs sampler approach. Increasing familiarity and experimentation with new Markov chain Monte Carlo methods for exploring and summarising posterior distributions in Bayesian statistics will lead to new insights in toxicokinetics and may provide a useful tool to handle datasets and problems as introduced in the present thesis.

Acknowledgements

The financial support of the Deutsche Forschungsgemeinschaft (SFB 475, "Reduction of complexity in multivariate data structures") is gratefully acknowledged.

References

Becka, M., Bolt, H. M. and Urfer, W. (1993). 'Statistical evaluation of toxicokinetic data'. *Environmetrics* **4**, 311-322.

Becka, M. (1994). Statistische Analyse kinetischer Daten in der toxikokinetischen Grundlagenforschung unter Verwendung von deterministischen Kompartiment-Modellen. Thesis, Department of Statistics, University of Dortmund.

Bolt, H.M. (1996). 'Quantification of endogenous carcinogens. The ethylene oxide paradox'. *Biochemical Pharmacology* **52**, 1-5.

Bolt, H.M. (1998). 'The Carcinogenic Risk of Ethene (Ethylene)'. *Toxicologic Pathology* **26**, 454-456.

Bolt, H.M. and Filser, J.G. (1987). 'Kinetics and disposition in toxicology. Example: Carcinogenic risk estimate for ethylene'. *Archives of Toxicology* **60**, 73-76.

Bolt, H.M., Filser, J.G. and Störmer, F. (1984). 'Inhalation pharmacokinetics based on gas uptake studies V. Comparative pharmacokinetics of ethylene and 1,3-butadiene in rats'. *Archives of Toxicology* **55**, 213-218.

Bolt, H.M., Leutbecher, M. and Golka, K. (1997). 'A note on the physiological background of the ethylene oxide adduct 7(2-hydroxyethyl)guanine in DNA from human blood'. *Archives of Toxicology* **71**, 719-721.

Büning, H. and Trenkler, G. (1994). *Nichtparameterische statistische Methoden*. Walter de Gruyter, Berlin, 243-247.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). 'Maximum likelihood from incomplete data via the EM algorithm'. *Journal of the Royal Statistical Society, Series B,* **39**, 1-38.

Filser, J.G. (1992). 'The closed chamber technique – uptake, endogenous production, excretion, steady-state kinetics and rates of metabolism of gases and vapours'. *Archives of Toxicology* **66**, 1-10.

Filser, J.G. and Bolt, H.M. (1983). 'Exhalation of ethylene oxide by rats exposed to ethylene'. *Mutation Research* **120**, 57-60.

Filser, J.G. and Bolt, H.M. (1984). 'Inhalation pharmacokinetics based on gas uptake studies. VI Comparative evaluation of ethylene oxide and butadiene monoxide as exhaled reactive metabolites of ethylene and 1,3-butadiene in rats'. *Archives of Toxicology* **55**, 219-223.

Filser, J.G., Denk, B., Törnqvist, M., Kessler, W. and Ehrenberg, L. (1992). 'Pharmacokinetics of ethylene in man; body burden with ethylene oxide and hydroxyethylation of hemoglobin due to endogenous and environmental ethylene'. *Archives of Toxicology* **66**, 157-163 [erratum: *Archives of Toxicology* **67**: 230, 1993].

Filser, J.G., Kreuzer, P.E., Greim, H. and Bolt, H.M. (1994). 'New scientific arguments for regulation of ethylene oxide residues in skin-care products'. *Archives of Toxicology* **68**, 401-405.

Gilks, W.R., Clayton, D.G., Spiegelhalter, N.G., Best, N.G., McNeil, A.J. and Kirby, A.J. (1993). 'Modelling complexity: applications of Gibbs sampling in medicine'. *Journal of the Royal Statistical Society, Series B* **55**, 39-52.

Hoel, D.G., Kaplan, N.L. and Anderson, M.W. (1983). 'Implication of nonlinear kinetics on risk estimation in carcinogenesis'. *Science* **219**, 1032-1037.

Hollander, M. and Wolfe, D.A. (1999). *Nonparametric Statistical Methods.* Wiley, New York, USA, 363-387, 413-414.

Hedrich, H. J. and Löscher, W. (1993). 'Impact of genetics on toxicity testing, introductory remarks'. *Journal of Experimental Animal Science* **35**, 197-199.

Kirkovski, L.I., Lermontov, S.A., Zavorin, S.I., Sukhozhenko, I.I., Zavelsky, V.I., Thier, R. and Bolt, H.M. (1998). 'Hydrolysis of genotoxic methyl-substituted oxiranes: experimental kinetic and semiempirical studies'. *Environmental Toxicology and Chemistry* **17**, 2141-2147.

Lindley, D.V. and Smith, A.F.M. (1972). 'Bayes estimates for the linear model' (with discussion). *Journal of the Royal Statistical Society, Series B,* **34**, 1-42.

Lovell, D. P. (1993). 'Impact of pharmacogenetics on toxicological studies. Statistical implications'. *Journal of Experimental Animal Science* **35**, 259-281.

Neave, H.R. and Worthington, P.L. (1992). *Distribution-free tests*. Routledge, London, UK.

Quinke, B., Selinski, S., Golka, K. and Blaszkewicz, M. (2000). 'Population toxicokinetics of ethylene: Calibration and preceding investigations'. *Technical Report 3/2000*, University of Dortmund. [http://www.statistik.uni-dortmund.de/sfb475/sfblit.htm]

Racine-Poon A. (1985). 'A Bayesian Approach to Nonlinear Random Effect Models'. *Biometrics* **41**, 1015-1023.

Racine-Poon A. and Smith A.F.M. (1990). 'Population models'. In: D.A. Berry (ed.): *Statistical Methodology in the Pharmaceutical Science*. Marcel Dekker, New York, USA, 139-162.

SAS STAT User´s Guide, Volume 2, Version 6 (1994). 4th edition, The SAS Institute, Cary, N.C., USA.

Schirm, F.A. (1999). *Schätzung von Populationsparametern eines toxikokinetischen Modells mit Hilfe von SAS*. Diploma Thesis, Department of Statistics, University of Dortmund.

Schirm, F.A. and Selinski, S. (2000). 'Interindividual and interoccasion variability of toxicokinetic parameters of uptake, exhalation, and metabolism of ethylene'. *Technical Report 7/2000*, University of Dortmund. [http://www.statistik.uni-dortmund.de/sfb475]

Selinski, S. (2000). 'Estimation of toxicokinetic population parameters in a four stage hierarchical model'. *Technical Report 1/2000*, University of Dortmund. [http://www.statistik.uni-dortmund.de/sfb475]

Selinski, S. (2001). *Bayes estimation of toxicokinetic parameters in a four-stage hierarchical model*. Thesis, Department of Statistics, University of Dortmund.

Selinski, S. and Becker, C. (2001). 'Outlier detection in experimental data using a modified Hampel identifier'. *Technical Report 40/2001*, University of Dortmund. [http://www.statistik.uni-dortmund.de/sfb475]

Selinski, S., Golka, K., Bolt, H.M. and Urfer, W. (2000). 'Estimation of toxicokinetic parameters in population models for inhalation studies with ethylene'. *Environmetrics* **11**, 479-495.

Selinski, S. and Urfer, W. (1998). 'Interindividual and interoccasion variability of toxicokinetic parameters in population models'. *Technical Report 38/1998*, University of Dortmund. [http://www.statistik.uni-dortmund.de/sfb475]

Thier, R. and Bolt, H.M. (2000). 'Carcinogenicity and genotoxicity of ethylene oxide: New aspects and recent advances'. *Critical Review in Toxicology* 30, 595-608.

Urfer, W. and Becka, M. (1996). 'Exploratory and model-based inference in toxicokinetics'. In *Statistics in Toxicology, ed. B. J. T. Morgan*. Oxford University Press, 198-216

Urfer, W., Mejza, S. and Hering, F. (1999). 'Quantitative trait loci mapping in plant genetics by alpha-design experiments and molecular genetic marker systems'. *Technical Report 34/1999*, University of Dortmund. (Available at: http//www.statistik.uni-dortmund.de/sfb475)

Wikle, C.K., Berliner, L.M. and Cressie, N. (1998). 'Hierarchical Bayesian space-time models'. *Environmental and Ecological Statistics* **5**, 117-154.