Chapter 2
Introduction: Heterogeneity in Medicine

Patients are not alike. In a therapeutic setting patients react differently to treatment regimes. Some patients do not respond to the treatment at all and others recover quickly under the treatment. This may depend on known factors such as genetic polymorphisms. For example, drugs known as beta blockers, which antagonize the beta-adrenergic receptor, are an important component of the treatment regimen for chronic heart failure (CHF). Genetic heterogeneity at the level of the beta-adrenergic receptor is thought to be a factor explaining the variable responses of CHF patients to beta blockade (DeGeorge and Koch 2007). Subsequent trials in search of personalized treatment of heart failure can take this genetic heterogeneity into account.

However, often the underlying mechanism that causes variability of treatment effects is not known; thus, the modest benefit observed in many clinical trials can be misleading because moderate average effects may be a mixture of substantial benefits for some, little, or no benefit for many and harm for a few. This is a case of unobserved heterogeneity of patients or populations since it is not possible to directly observe to which subpopulation a patient belongs. Likewise in this case, the underlying covariate which causes the variability in treatment response is not known.

From a statistical point of view this warrants a method which identifies the presence of unobserved heterogeneity and in a second step enables us to take known covariates into account to explain some of the heterogeneity. This leads to subject-specific or random effects models. In this type of modeling one is interested in investigating the individual departures from the mean response in order to analyze the results at the level of the individual patient. In these models the variability between individuals is assumed to follow a random distribution. In principle this can be any proper distribution.

This work focuses on a discrete unobserved distribution for the variability between individuals, which leads to finite mixture models.

Coming back to the possible outcome of a clinical trial with a modest average effect, there are three latent subpopulations: one with substantial benefit, one with little benefit, and one subpopulation which is harmed. The idea of a finite mixture model is now to estimate the proportion of these respective subpopulations and the

P. Schlattmann, *Medical Applications of Finite Mixture Models*,
Statistics for Biology and Health, DOI: 10.1007/978-3-540-68651-4_2,

corresponding mean treatment effect in these groups. This approach resembles an analysis of variance, albeit the membership of an individual to a certain subpopulation is not known. As an example consider the outcome of a fictitious clinical trial dealing with asthma. Here improvement of the forced expiratory volume (FEV1) may be a dependent variable. Suppose that the average improvement of a new drug was 221 ml in comparison with a placebo. Let us assume that the population of these patients consists of subpopulations due to a genetic polymorphism which is not yet known. For simplicity we postulate that only heterozygote alleles and not homozygote alleles are of importance, which implies that only two subpopulations exist. In this mind experiment we assume that the majority of patients do not benefit greatly from the drug, that is 70% of the patients show an improvement of FEV1 of only 80 ml compared with the placebo. On the other hand, 30% of the patients benefit from an improvement of 550 ml.

In a first model we assume that the data follow a normal distribution, $N(\mu, \sigma^2)$, where $\mu = 221$ ml denotes the mean and $\sigma^2 = 40,000$ the variance. A second model takes the presumed heterogeneity of populations into account. It consists of a weighted sum of normal distributions with

$$y \sim 0.7 \times N(80, 22, 500) + 0.3 \times N(550, 22, 500), \tag{2.1}$$

where the common standard deviation is assumed to be 150 ml. These two models are depicted graphically in Fig. 2.1.

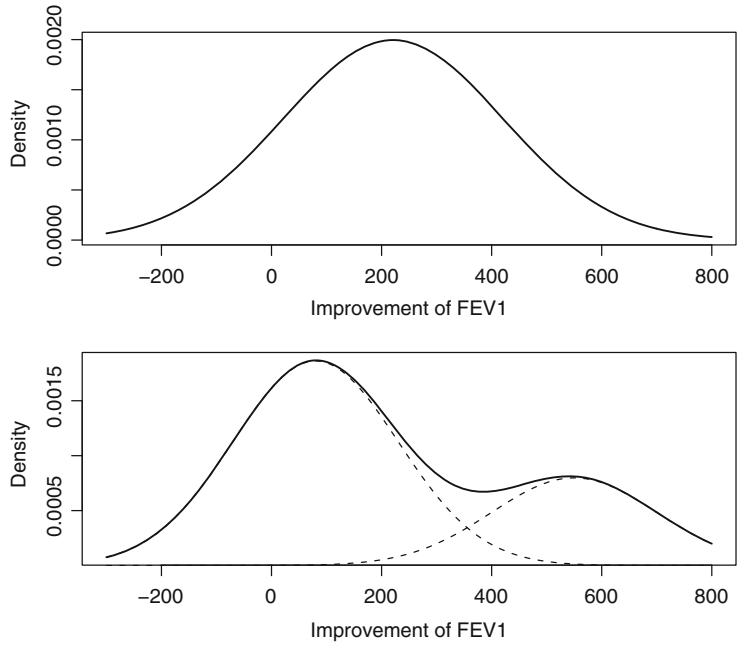

Fig. 2.1 Homogenous model with mean treatment effect $\mu = 221$ ml and a model with two subpopulations. Seventy percent of patients with little benefit ($\mu_1 = 80$ ml) and 30% of patients with considerable benefit ($\mu_2 = 550$ ml). *FEV1* forced expiratory volume

Table 2.1 Simulated data from a mixture of normal distributions

528.14	−49.63	283.65	152.99	472.42	365.29
672.86	167.65	649.12	338.38	612.13	115.80

Table 2.2 Simulated data from a two-component mixture of normal distributions with indicator variables z with regard to their component membership

Data y	z_1	z_2
528.14	0	1
−49.63	1	0
283.65	1	0
152.99	1	0
472.42	0	1
365.29	0	1
672.86	0	1
167.65	0	1
649.12	0	1
338.38	0	1
612.13	0	1
115.80	1	0

Table 2.1 shows simulated data from the second model. They correspond to the data directly observable by an investigator analyzing the results of the clinical trial. Mean and corresponding proportions of the respective subpopulations are not known. Likewise, for an individual observation it is not known to which subpopulation it belongs.

Now, if the component membership of an individual were known, estimation of the corresponding subpopulation's mean and the corresponding mixture proportion would be easy: To describe component membership, an indicator variable Z is introduced. If an observation belongs to the first component, the indicator variable z_1 will take the value 1 and 0 otherwise. For the second subpopulation the indicator variable z_2 is defined similarly. Table 2.2 shows the known indicator variables of the component membership of a certain individual.

Using the known indicator variables in Table 2.2, it turns out that the proportion of the first component equals the proportion of "ones" of the first component z_1, given by

$$\hat{p}_1 = \frac{4}{12} = 0.333. \tag{2.2}$$

The proportion of the second component z_2 is simply

$$\hat{p}_2 = \frac{8}{12} = 0.667 \quad \text{or} \quad \hat{p}_2 = 1 - \hat{p}_1. \tag{2.3}$$

In the same manner each subpopulation's mean can be computed. The mean μ_1 of the first component is given by

$$\hat{\mu}_1 = \frac{-49.63 + 283.65 + 152.99 + 115.80}{4} = 125.71. \tag{2.4}$$

Similarly, the mean μ_2 of the second component can be computed as

$$\hat{\mu}_2 = \frac{528.14 + 472.42 + \cdots + 338.38 + 612.13}{8} = 475.75. \tag{2.5}$$

Unfortunately, in general, component membership is an unobservable variable. In this case, a frequently used algorithm for finite mixture models (see Sect. 4.4.3), proceeds as follows. First, determine the number of subpopulations k. Then, provide a first guess for the relative frequencies $p_1, \ldots, p_k$ of each component and the associated mean values $\mu_1, \ldots, \mu_k$. Applying Bayes's theorem (see (4.38) on page 75 for details) this information allows one to calculate e_{ij}, which denotes the probability that the ith observation belongs to the jth component.

By replacing the unknown indicator variables with the probabilities of component membership, one obtains new estimates of the mixing proportions p_j and population means μ_j. This procedure is repeated until a convergence criterion is met. This algorithm is known as the expectation maxmization. A formal derivation and description of algorithms for finite mixture models is developed in Chap. 4.

In contrast to Fig. 2.1, mixtures of normal densities with two components are not necessarily bimodal. This is shown in Fig. 2.2. The shape of mixture densities of

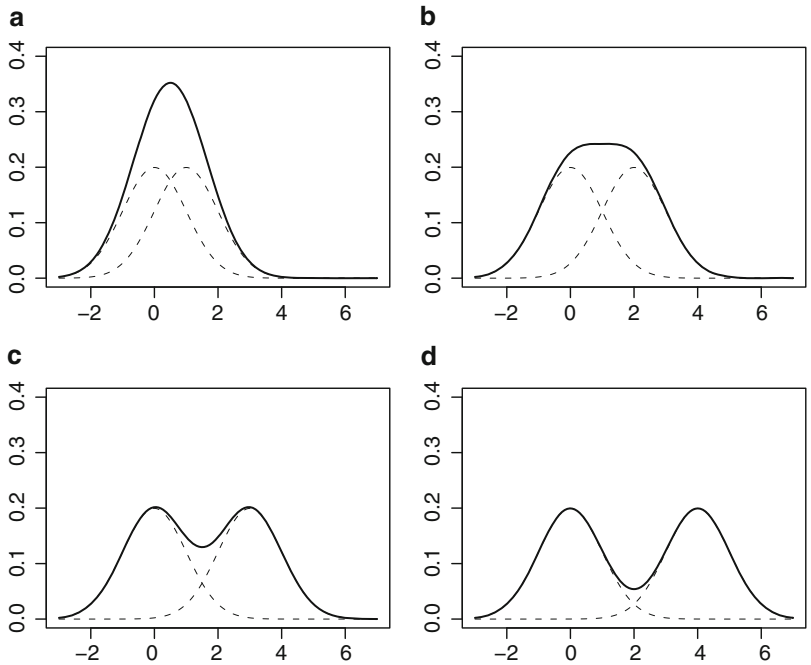

Fig. 2.2 Mixture densities with two univariate normal distributions with $p_1 = p_2 = 0.5$, $\mu_1 = 0$, and $\sigma^2 = 1$ in the cases **a** $\mu_2 = 1$, **b** $\mu_2 = 2$, **c** $\mu_2 = 3$, and **d** $\mu_2 = 4$

two univariate normal densities depends on how far the individual means of the two distributions are apart. Figure 2.2 shows mixtures with $\mu_1 = 0$, $p_1 = p_2 = 0.5$, and variance $\sigma_2 = 1$. If the difference $\Delta = |\mu_1 - \mu_2|$ is 1 (see Fig. 2.2a) the shape of the mixed distribution is unimodal. The same applies if the difference is 2. For $\Delta = 3$ and $\Delta = 4$ the resulting mixture distributions are bimodal.

Following this presentation of the general ideas of finite mixture models, Sect. 2.1 applies this model to data from an epidemiologic study.

2.1 Example: Plasma Concentration of Beta-Carotene

2.1.1 Identification of a Latent Structure

High intakes of fruits and vegetables, or high circulating levels of their biomarkers (carotenoids, vitamins C and E), have been associated with a relatively low incidence of cardiovascular disease, cataracts, and cancer. A high fruit and vegetable diet increases antioxidant concentrations in blood and body tissues, and potentially protects against oxidative damage to cells and tissues. This observation led to the initiation of randomized clinical trials focusing on subjects with a high risk of cancer. One of these trials is the beta-Carotene and Retinol Efficacy Trial (CARET) initiated by Goodman et al. (1993). This trial tested the effect of daily beta-carotene (30 mg) and retinyl palmitate (25,000 IU) intake on the incidence of lung cancer, other cancers, and death in 18,314 participants who were at high risk for lung cancer because of a history of smoking or asbestos exposure. CARET was stopped ahead of schedule in January 1996 because participants who were randomly assigned to receive the active intervention were found to have a 28% increase in incidence of lung cancer, a 17% increase in incidence of death, and a higher rate of cardiovascular disease mortality compared with participants in the placebo group (Goodman et al. 2004).

In a systematic review summarizing the evidence from controlled clinical trials Caraballoso et al. (2003) concluded that there is currently no evidence to support recommending vitamins such as alpha-tocopherol, beta-carotene, or retinol, alone or in combination, to prevent lung cancer. Likewise, the risk of development of nonmelanoma skin cancer was not found to be related to serum levels of any of the carotenoids measured in a study performed by Dorgan et al. (2004). Even worse, a recent meta-analysis of randomized trials (Bjelakovic et al. 2007) found increased mortality of patients supplemented with antioxidants.

Despite these disappointing results with regard to supplementation of antioxidants, the investigation of the determinants of plasma concentrations of micronutrients is still an ongoing area of research (Goodman et al. 1996; Margetts and Jackson 1996; Lagiou et al. 2003; Faure et al. 2006). A common finding in these studies is a negative association between cigarette smoking and plasma levels of beta-carotene. Many of the investigators mentioned before used a linear regression model to investigate the association between factors such as smoking, age, or gender and plasma

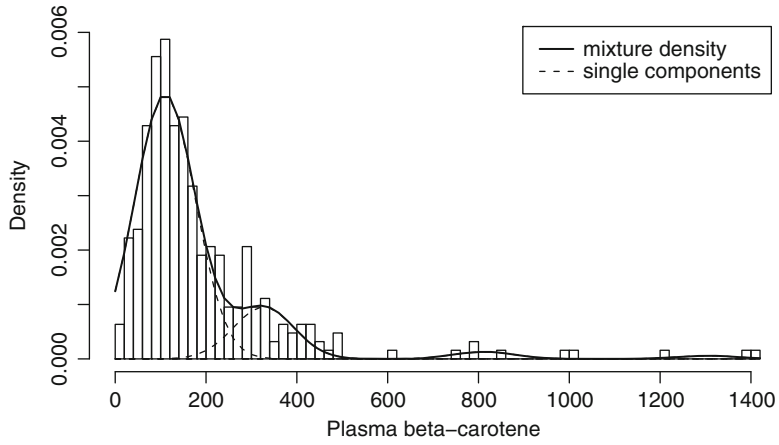

Fig. 2.3 Beta-Carotene plasma concentrations together with a four-component mixture model and single mixture components

levels of antioxidants. In our example, we use the data on beta-carotene plasma levels from Nierenberg et al. (1989) and Stukel (2008)[1] to address the following questions: Are there latent subgroups present in this data set and, if so, can this heterogeneity be explained by covariates such as age, gender, or smoking?

Figure 2.3 shows that there is wide variability in subjects. Especially the assumption of a normal distribution for these data seems not to be appropriate. For that reason Nierenberg et al. (1989) used a log transformation of the data to find determinants of plasma beta-carotene concentrations. But the multimodal shape of the histogram suggests that the data were sampled from a population which consists of several homogenous subpopulations.

Thus, the first step of an alternative analysis of these data tries to identify the latent structure. Using the methods described in Chap. 4, Sect. 4.4.5 leads to a solution which identifies four latent subpopulations. The first group has a mean concentration of beta-carotene of 126.05 ng ml^{-1} and constitutes 80.2% of the whole sample. The next latent population has a mean concentration of 350.83 ng ml^{-1} and forms another 16.8% of the total sample. The third subgroup has a mean concentration of 854.55 ng ml^{-1} and constitutes 2% of the population. Finally, the last subgroup has a mean concentration of 1,339.32 ng ml^{-1} and forms 1% of the overall population. The variance is assumed to be the same for each subpopulation and is 4,371.2. Hence, instead of a single normal distribution, these data can be described by a weighted sum of normal densities, that is,

$$\begin{aligned} f(x) = {} & 0.802 \times N(126.05, 4371.2) + 0.168 \times N(350.83, 4371.2) + 0.02 \\ & \times N(854.55, 4371.2) + 0.01 \times N(1339.33, 4371.2). \end{aligned} \tag{2.6}$$

[1] The use of the data for this book and the accompanying R package is kindly permitted by Therese Stukel, Dartmouth Hitchcock Medical Center, USA.

Looking at the fitted line in Fig. 2.3, we see that this finite mixture model provides an acceptable fit to the data.

2.1.2 Including Covariates

After identification of latent subpopulations, the question arises whether this latent structure can be explained by known covariates. The corresponding theory of covariate-adjusted finite mixture models may be found in Sect. 4.6.

In the simplest case this leads to a model with random intercepts. This implies that there are several regression models for each subpopulation which have the same slope but different intercepts. For example, consider the association of the amount of beta-carotene in the patient's diet with the beta-carotene plasma level of that individual. This covariate is labeled *diet*. In the Sect. 2.1.1 four subpopulations were identified. Now if the effect of diet is assumed to be the same in each subpopulation, this leads to the following four regression equations:

$$\begin{aligned}
\hat{\mu}_{i1} &= 109.80 + 0.0087 \times \text{diet}_i \,, \\
\hat{\mu}_{i2} &= 327.23 + 0.0087 \times \text{diet}_i \,, \\
\hat{\mu}_{i3} &= 812.30 + 0.0087 \times \text{diet}_i \,, \\
\hat{\mu}_{i4} &= 1303.52 + 0.0087 \times \text{diet}_i \,.
\end{aligned} \tag{2.7}$$

Since the effect of diet is assumed to be identical in each subpopulation, it is often called a fixed effect. This model is depicted on the left-hand side of Fig. 2.4. Here, the first component with the lowest intercept constitutes 81% of the total sample. The second component forms 16% of the population. As before, in the model without covariates the third component contributes 2% and the fourth component 1%. In terms of interpretation, the majority of patients have rather low beta-carotene plasma levels, 16% have intermediate concentrations, and only 3% have high concentrations.

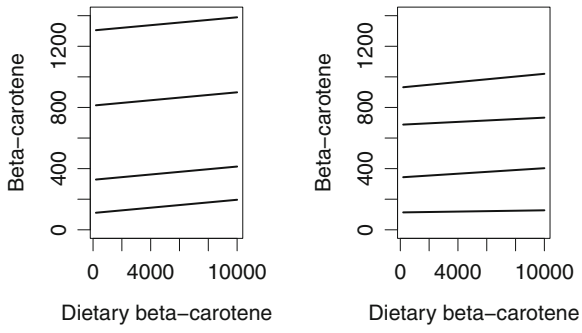

Fig. 2.4 Beta-Carotene data: finite mixture models adjusted for covariates. *Left*: A model with random intercepts. *Right*: A model with random slopes and intercepts

In a more complex model the effect of the covariate may differ in each subpopulation as well. This leads to a regression model with different intercept and slope for each subpopulation:

$$\begin{aligned}\hat{\mu}_{i1} &= 113.56 + 0.0014 \times \text{diet}_i\,,\\ \hat{\mu}_{i2} &= 342.69 + 0.0062 \times \text{diet}_i\,,\\ \hat{\mu}_{i3} &= 687.04 + 0.0046 \times \text{diet}_i\,,\\ \hat{\mu}_{i4} &= 930.32 + 0.0089 \times \text{diet}_i\,. \end{aligned} \tag{2.8}$$

This is depicted on the right-hand side of Fig. 2.4. Now the effect of diet varies between subpopulations and is assumed to follow a random distribution; this is often called a random effect. The corresponding mixing proportions are given by $\hat{p}_1 = 0.8$, $\hat{p}_2 = 0.17$, $\hat{p}_2 = 0.02$, and $\hat{p}_4 = 0.01$. According to this model, 80% of the individuals have a small intercept of 113.56 with a small effect of dietary beta-carotene intake. Then there is an intermediate group with an intercept of 342.69 and a relatively large effect of dietary beta-carotene intake equal to 0.0062. Finally, there are two small subpopulations with large intercepts and a intermediate effect of the covariate dietary beta-carotene intake. After these preliminary considerations other covariates can be included into the model as well. Nierenberg et al. (1989) found a positive association between beta-carotene levels, dietary carotene, and female gender. Cigarette smoking and body mass index were negatively related to beta-carotene levels. Use of vitamins was also positively associated with beta-carotene plasma levels, whereas age was not associated with beta-carotene levels to a statistically significant extent.

Considering a covariate-adjusted finite mixture model, one obtains the results shown in Table 2.3. According to this model there are four subpopulations. The corresponding mixture proportions are given by $\hat{p}_1 = 0.8$, $\hat{p}_2 = 0.17$, $\hat{p}_2 = 0.02$, and $\hat{p}_4 = 0.01$. Again, the effect of dietary beta-carotene is different in these subpopulations. In the first subpopulation, constituting 80%, the effect is rather small, whereas

Table 2.3 Beta-Carotene data: finite mixture model adjusted for covariates gender, body mass index (*BMI*), smoking status (current smoker vs. never; former smoker vs. never), and dietary beta-carotene

Component	Weight $\hat{p}_j$	Estimate	Standard error	t	$\Pr(>\|t\|)$
Intercept 1	0.80	187.472	13.903	13.484	<0.001
Intercept 2	0.17	405.898	16.529	24.557	<0.001
Intercept 3	0.02	768.694	31.59	24.333	<0.001
Intercept 4	0.01	1,005.866	54.609	18.419	<0.001
BMI		−3.637	0.405	−8.980	<0.001
Current vs. never		−31.82	7.451	−4.271	<0.001
Former vs. never		−19.935	5.318	−3.749	<0.001
Female vs. male		32.950	7.105	4.638	<0.001
Diet 1		0.002	0.003	0.667	0.414
Diet 2		0.089	0.002	44.5	<0.001
Diet 3		0.042	0.009	4.667	<0.001
Diet 4		0.084	0.013	6.462	<0.001

in the fourth subpopulation the effect of dietary beta-carotene is much larger. Other covariates did not turn out to behave differently in these subpopulations. In accordance with Nierenberg et al. (1989) there is a negative association between the body mass index, smoking status, and male gender. On the basis of this model, vitamin use and age are not associated with beta-carotene levels to a statistically significant extent. In conclusion, this type of model has at least two advantages. First, we could identify subpopulations who react differently to dietary beta-carotene. This might be a starting point for further investigations. Especially of interest would be the question whether the recent adverse findings of supplementation of antioxidants apply to all patients similarly. Second, this type of modeling is one way to handle the problem that the data apparently do not follow a normal distribution. In contrast to, e.g., a logarithmic transformation of the data, the interpretation of the results remains on the original scale of the observations.

2.2 Computation

The previous models may be fit using the R package CAMAN. The first step in using the program is to load the package and the data set. This is done typing

```
> library(CAMAN)
> data(betaplasma)
```

The next step is to identify the latent structure of the data. To obtain a first impression of the potential number of subpopulations in our data the vertex exchange method (VEM) algorithm is applied. This algorithm is based on a fixed grid of potential means of subpopulations. The default in *mixalg.VEM* calculates the minimum and the maximum of the data and constructs a grid with $k = 50$ equidistant grid points. Then, the algorithm finds the corresponding population proportions that maximize the likelihood function. A detailed description of the algorithm may be found in Sect. 4.3.2.

The function is called with

```
> beta0<-mixalg.VEM(obs="betacaro",data=betaplasma,
startk=50,family="gaussian")
```

Here the call takes as arguments the dependent variable *betacaro*, which is the plasma level of beta-carotene. The number of grid points is set to $k = 50$. By default the mixing density is set to be the normal distribution. Also as a default the variance is set equal to the empirical variance. Typing *beta0* provides the following output:

```
> beta0

Computer Assisted Mixture Analysis (VEM):

Data consist of 315 observations (rows) four grid
points with positive support
```

```
 p             parameter
0.973285469  173.2653
0.017760636  895.2041
0.001319593 1299.4898
0.007634302 1328.3673

Log-Likelihood:   -2029.80100  BIC:  4629.10600
```

This result suggests that the data could be described by a mixture model with four components. In the next step we apply the function *mixcov*, which allows the inclusion of covariates into the model. The call is given by

```
> beta1<-mixcov(c(dep="betacaro"),fixed=c("1"),
random=c(""),data=betaplasma,k=4,family="gaussian")
```

This gives the (shortened) result

```
> beta1

Computer Assisted Mixture Analysis with covariates:

Data consist of 315 observations (rows).


mixing weights:
    comp. 1     comp. 2     comp. 3     comp. 4
     0.8026       0.168      0.0199      0.0095

Coefficients :
      Z1       Z2       Z3       Z4
  126.0613   350.826  854.549 1339.32978

Log-Likelihood: -1927.117     BIC: 3894.50
```

In the next step the covariate *betadiet* is included as a fixed effect into the model.

```
beta2<-mixcov(c(dep="betacaro"),fixed=c("betadiet"),
random=c(""),data=betaplasma,k=4,family="gaussian")
```

This results in the following output:

```
> beta2

mixing weights:
    comp. 1     comp. 2     comp. 3     comp. 4
0.811028183 0.157662064 0.021785704 0.009524049
```

```
Coefficients :
    Z1       Z2       Z3       Z4 betadiet
 109.855  327.232  812.343  1303.201   0.009

Log-Likelihood: -1924.664     BIC: 3895.349
```

Inclusion of the covariate leads to an improvement of the likelihood. Thus, not surprisingly, we deduce an association between dietary beta-carotene and plasma levels of beta-carotene. Next, we fit a model with the covariate *betadiet* as a random effect.

```
beta3<-mixcov(c(dep="betacaro"),fixed=c("1"),random=
c("betadiet"),data=betaplasma,k=4,family="gaussian")
```

This leads to the result

```
> beta3

mixing weights:
   comp. 1    comp. 2    comp. 3    comp. 4
0.80164319 0.16663101 0.01913382 0.01259198

Coefficients :
     Z1          Z2          Z3        Z4
    113.549     342.815     687.089     930.381

    Z1:betadiet Z2:betadiet Z3:betadiet Z4:betadiet
    0.001       0.006       0.046       0.090

Log-Likelihood: -1919.450     BIC: 3902.179
```

Comparing the log likelihoods of these models, we find again an improvement when we allow varying effects of the covariate in the respective subpopulations.

More covariates can be included into the model, for example, we might be interested in the effect of smoking status. Here, we use the treatment contrasts *current vs. never* and *former vs. never* The data set contains the variable *smokestat*, which is defined as a factor variable. By default R uses treatment contrasts for factor variables Now, the model is obtained with

```
beta4<-mixcov(c(dep="betacaro"),fixed=c("smokestat"),
random=c("betadiet"),data=betaplasma,k=4,
family="gaussian")
```

The result is given by

```
> beta4

mixing weights:
   comp. 1    comp. 2    comp. 3    comp. 4
0.81527849 0.15298759 0.01911870 0.01261522
```

```
Coefficients :
             Z1                 Z2          Z3          Z4
        125.882           354.613     704.594     959.963

smokestatFormer smokestatCurrent
        -20.667           -30.533

Z1:betadiet     Z2:betadiet    Z3:betadiet     Z4:betadiet
0.007           0.002          0.045           0.084

Log-Likelihood: -1915.798     BIC: 3906.379
```

2.3 Example: Analysis of Heterogeneity in Drug Development

2.3.1 Basic Pharmacokinetic Concepts

Pharmacokinetics investigates the absorption and disposition of drugs. Disposition is further subdivided into the investigation of distribution, metabolism, and excretion of a drug. Thus, pharmacokinetics is sometimes referred to as ADME (absorption, distribution, metabolism, excretion). A general introduction to the field of pharmacokinetics may be found in the books by Winter (2004), Rowland and Tozer (2005), and Tozer and Rowland (2006). The absorption characteristics of a drug depend on the properties of the chemical compound as well as on the route of administration and the exact formulation used. Obviously, this is not of interest for drugs which are administered by infusion or intravenous injection. On the other hand, absorption characteristics play an important role in the kinetics of orally administered drugs. The rate and the extent of absorption are important pharmacokinetic quantities for orally administered drugs.

Once in circulation, a drug is distributed throughout the body. Owing to differing characteristics of the various tissue types, this distribution is not uniform. Rather, areas of higher or lower concentration can be observed. Again, the chemical characteristics of the drug, and how it interacts with its surroundings at the molecular level, are determinants of this behavior.

Metabolism and excretion are the two ways in which a drug is removed from the body. Metabolism, which is also called "biotransformation," takes place mainly in the liver. The liver contains a host of enzymes, the cytochrome P450 (CYP) family, whose function is to dispose of chemicals that have entered the body (Pelkonen and Breimer 1994). These enzymes use iron to oxidize substances, often as part of the body's strategy to dispose of potentially harmful substances by making them more water-soluble. Bertz and Granneman (1997) found that 56% of 315 drugs were primarily cleared by CYP enzymes.

The products of biotransformation are called "metabolites." For some drugs, it is a metabolite rather than the parent compound that is the active substance. This applies, for example, to Dipyrone, an analgesic, antipyretic, and anti-inflammatory drug which is studied in Sect. 4.7.2.

2.3.2 Pharmacokinetic Parameters

The study of pharmacokinetics is a central part of drug development. For the analysis of pharmacokinetic data two different approaches are available. The first one is the so-called noncompartmental approach which relies on parameters such as maximum concentration $C_{\max}$, the time of maximum concentration $T_{\max}$, and the area under the time–concentration curve. The second approach is given by so-called compartment models. Compartment models describe the flow of a substance (chemicals, drugs, information) between the components of a larger system. In general, there are m components (the compartments) which may be linked to each other in any way. The flow rates or exchange constants between linked compartments are generally specified, and each compartment may also have output and input to and from the outside world. Compartments are described by differential equations. An overview of the use of compartment models in pharmacokinetics may be found in the books by Notari (1975) and Gibaldi and Perrier (1982). Basic pharmacokinetic parameters are defined and explained next.

Absorption Rate

In pharmacokinetics, absorption is the movement of a drug into the bloodstream. Absorption is often assumed to be a first-order process (Notari 1975), meaning that the rate of transfer is proportional to the amount to be transferred. In this case, the absorption rate (k_a) represents the constant of proportionality.

Elimination Rate

Similar to the absorption rate, the elimination rate (k_e) is used to refer to the proportionality constant in a first-order elimination process. Elimination refers to both biotransformation and excretion, and many processes can participate in the elimination of a drug. Here, k_e refers to the total elimination from all these processes. Elimination is usually well estimated by a first-order process. This implies that a constant fraction of the drug in the body is eliminated per unit time; thus, the rate of elimination is proportional to the amount of drug in the body. Even though some elimination processes may in fact be capacity-limited (e.g., in cases where enzymes are involved), the typically low concentrations associated with therapeutic doses imply that this usually does not affect first-order behavior (Gibaldi and Perrier 1982).

Apparent Volume of Distribution

The apparent volume of distribution (V) of a drug is defined by

$$V = \frac{x}{c}, \tag{2.9}$$

where x is the amount of drug in the body and c denotes the concentration in blood or plasma. V is not a physical volume because the distribution of the drug in the body is nonuniform. It will not be lower than the blood or the plasma volume but for some drugs it can be much larger than the body volume. The volume of distribution is a mathematical factor relating the amount of drug in the body and the concentration of drug in the measured compartment, usually plasma.

Clearance

Clearance (Cl) can be interpreted as the volume from which all drug molecules are eliminated during a certain time span. This parameter is defined by

$$\mathrm{Cl} = V k_{\mathrm{e}}. \tag{2.10}$$

Thus, clearance relates the volume of distribution and the elimination rate, and is a measure for the rate of elimination of a drug (Cawello 1999). The clearance is often considered to be the most important pharmacokinetic parameter (Gibaldi and Perrier 1982). Individual clearances can be defined for each elimination organ. Here only the total body clearance will be considered, i.e., the sum of all clearances.

2.3.3 First-Order Compartment Models

One of the simplest compartment models is a first-order oral compartment model. The principal idea is shown graphically in Fig. 2.5. This model can be obtained as a solution to a differential equation and describes the plasma concentration $c(t)$ at time t. The corresponding nonlinear regression model is given in (2.11):

$$c(t) = \frac{D k_{\mathrm{a}} k_{\mathrm{e}}}{\mathrm{Cl}(k_{\mathrm{a}} - k_{\mathrm{e}})} (\mathrm{e}^{-k_{\mathrm{e}} t} - \mathrm{e}^{-k_{\mathrm{a}} t}). \tag{2.11}$$

In this equation k_{a} and k_{e} represent the constants of absorption and elimination, respectively, whereas Cl denotes the clearance and D denotes the dose given. This model implies that at the beginning the concentration–time curve is governed by the absorption constant k_{a} and later this curve is determined by the elimination constant k_{e}. This is shown schematically in Fig. 2.6. Equation (2.11) forms a nonlinear regression problem. In general there exist no closed-form solutions for nonlinear regression models. A solution has to be found by numerical techniques such as

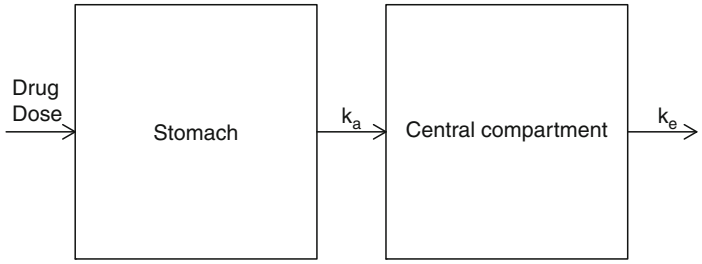

Fig. 2.5 The one-compartment model. The drug is absorbed from the gastrointestinal tract and then eliminated. k_a and k_e represent the rates of absorption and elimination, respectively

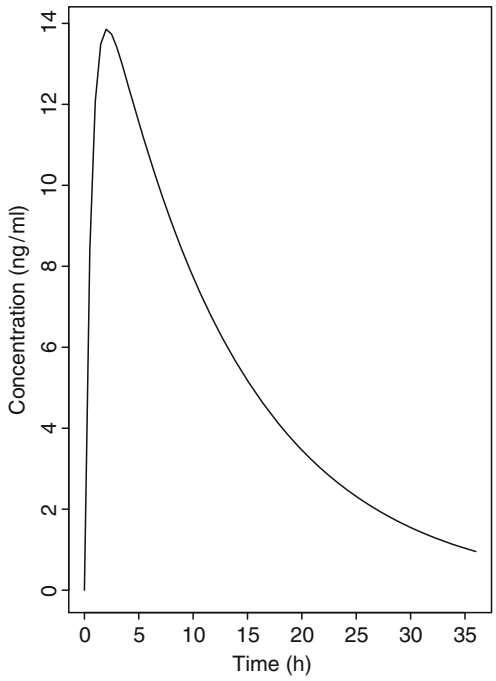

Fig. 2.6 Concentration–time curve based on an oral one-compartment model

the Gauss–Newton algorithm. For details on nonlinear regression problems see the monograph by Bates and Watts (1988). Many of the algorithms proposed and described in that monograph are part of the R package *nlme* (Pinheiro and Bates 2000).

2.3.4 Population Pharmacokinetics

The foundations of population pharmacokinetic modeling were laid in the 1970s by Sheiner et al. (1972, 1977). This approach aims to model the relationship between

physiologic function (normal and disease altered) and pharmacokinetics while taking into account the interindividual variability in these relationships (Steimer et al. 1994). Thus, again the population pharmacokinetic approach tries to take heterogeneity of individuals into account. Variability of the pharmacokinetic response between individuals may be caused by differences in, e.g., absorption or elimination. These differences may be a result of genetic polymorphisms as mentioned in the introductory example. These differences may also be due to factors such as age, gender, or reduced renal or hepatic function. Often these reasons are not known. Hence, we search for a model which allows us to handle variability between individuals due to unobserved covariates.

As a result, the population pharmacokinetic approach has been gaining popularity. Meanwhile it is an important part of the drug development process. In 1999, the US FDA published a guidance for industry regarding population pharmacokinetics. It underlines the interest in population pharmacokinetics within the pharmaceutical industry (U.S. Department of Health and Human Services: Food and Drug Administration 1999). They conclude:

> Use of the population pharmacokinetic approach can help increase understanding of the quantitative relationships among drug input patterns, patient characteristics, and drug disposition. . . .

Another major asset of the method may be seen in the fact that sparse routinely collected data may be used. In the 1970s, when Sheiner et al. laid the foundations of population pharmacokinetic analyses, they showed that with data collected as part of routine patient care such modeling can estimate the average values of pharmacokinetic parameters and the interindividual variances of those parameters in a patient population. With such sparsely sampled data from patients receiving digoxin, their method produced estimates that were similar to published values derived with traditional methods (Sheiner et al. 1975). In the following section the ideas of population pharmacokinetic modeling are introduced using a data set on the pharmacokinetics of theophylline.

2.3.5 Theophylline Pharmacokinetics

Theophylline, also known as dimethylxanthine, is a drug used in therapy of respiratory diseases such as chronic obstructive pulmonary disease and asthma. Initial metabolism of theophylline is primarily performed by the hepatic CYP enzymes CYP1A2 and CYP2E1 (Tjia et al. 1996; Yoon et al. 2006). These data are taken from a phase I clinical study (Boeckmann et al. 1994). Figure 2.7 shows the concentration–time curve of 12 individuals whose theophylline concentrations were measured at 11 points in time after oral administration of one dose. Looking at these individual time–concentration curves, one sees a first-order oral compartment model seems appropriate. The individual dose given depends on the body weight of the subject.

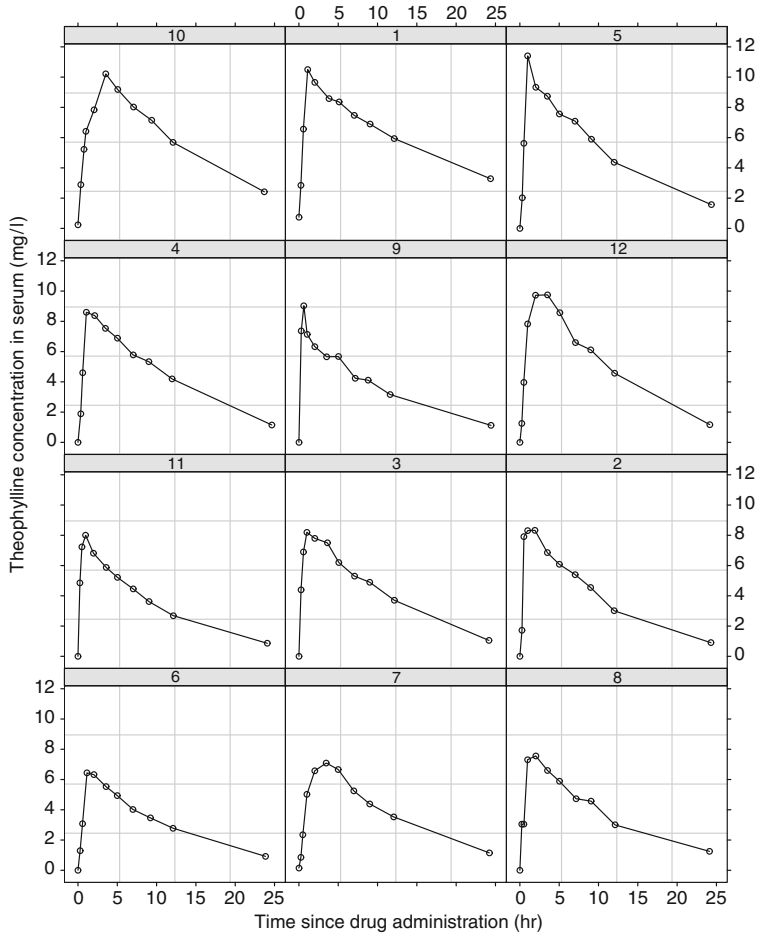

Fig. 2.7 Concentration–time curves of 12 subjects receiving an orally administered dose of theophylline

Table 2.4 shows the weight of the subjects, the dose given, and the noncompartmental pharmacokinetic parameters $T_{\max}$ and $C_{\max}$.

The simplest possible pharmacokinetic model is given by a so-called pooled model which assumes for all individuals the same pharmacokinetic model with identical parameters for absorption rate, elimination rate, and clearance. The corresponding result is shown in Fig. 2.8. The concentration–time curves of these individuals differ since they have received different doses; however, the plots indicate substantial interindividual variability. It becomes clear that the assumption of homogeneity is too strong and that this model does not provide a satisfactory fit to the data. Since this data set consists of 11 concentration measurements for each of the 12 subjects, an individual model for each subject could be calculated alternatively.

Table 2.4 Subject characteristics and pharmacokinetic parameters T_{max} and C_{max}

Subject	Weight (kg)	Dose (mg)	C_{max} (ng ml^{-1})	T_{max} (h)
6	80.0	4.00	6.44	1.15
7	64.6	4.95	7.09	3.48
8	70.5	4.53	7.56	1.02
11	65.0	4.92	8.00	0.98
3	70.5	4.53	8.20	1.02
2	72.4	4.40	8.33	1.92
4	72.7	4.40	8.60	1.07
9	86.4	3.10	9.03	0.63
12	60.5	5.30	9.75	3.52
10	58.2	5.50	10.21	3.55
1	79.6	4.02	10.50	1.12
5	54.6	5.86	11.40	1.00

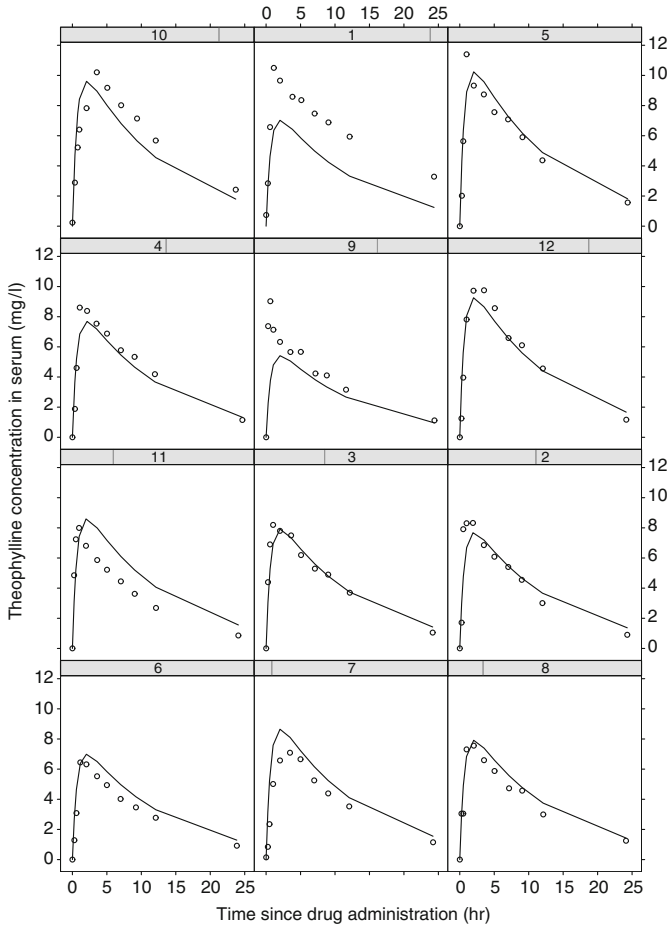

Fig. 2.8 Pooled fixed effects model for concentration–time curves of 12 subjects receiving an orally administered dose of theophylline

However, fitting a model to each individual implies fitting $12 \times 3 = 36$ parameters for the whole data set. As a result, this approach does not provide substantial data reduction. Thus, neither a pooled model nor individual models are particularly useful. One model for all individuals is too strict and one model for each individual does not provide sufficient data reduction. Additionally, many population pharmacokinetic studies do not provide sufficient data to be able to fit a model to each individual.

Thus, alternatively a random effects model can be fit to the data. Here again, variability between subjects can be modeled using a finite mixture model. In general, subpopulations with different absorption, elimination and clearance are considered. More precisely, there might be a subpopulation which has a large absorption constant but a low clearance. This would result in high, long-lasting serum levels of the drug. On the other hand, there might be a subpopulation with a small absorption constant but a high clearance. This would result in short-lasting, lower concentrations of the drug. For each of these subpopulations the elimination constant is assumed to be identical. In other words, absorption and clearance are assumed to vary between individuals and the elimination constant is assumed to be fixed. Technically speaking, absorption and clearance are assumed to follow a random distribution: They are random effects in the model.

The assumption of random effects for absorption and clearance but a fixed effect for elimination is motivated by Fig. 2.9. This figure shows the absorption rate, elimination rate, and clearance on a log scale for each individual together with a 95% confidence interval. It becomes clear that there is considerable variability between individuals for absorption and clearance, but less so for elimination. With these considerations a finite mixture model is fit to the data. The results are shown in Table 2.5. It turns out that there are three subpopulations. One subpopulation forms 34.58% of all individuals with a small absorption constant and relatively large clearance.

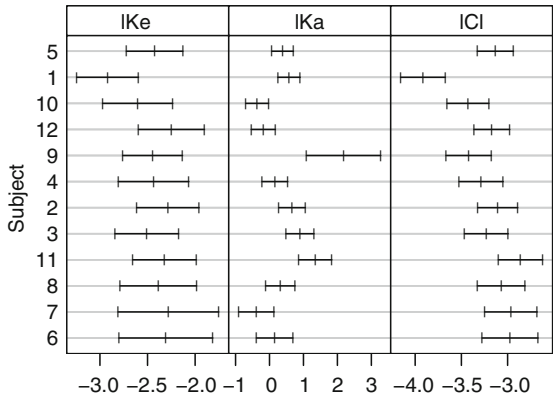

Fig. 2.9 Confidence intervals of individual one-compartment models of 12 subjects receiving an orally administered dose of theophylline, where *lKe* denotes the log of the elimination constant, *lKa* denotes the log of the absorption constant, and *lCl* denotes the log of the clearance

Table 2.5 Parameter estimates for the three-component mixture model fit to the theophylline data. Residual error is expressed as the standard deviation

Parameter	Group 1 (34.58%)	Group 2 (48.75%)	Group 3 (16.67%)
$\hat{k}_e$	0.082	0.082	0.082
$\hat{k}_a$	0.883	1.9332	2.409
$\hat{Cl}$	0.038	0.042	0.028
Residual error		1.0247	

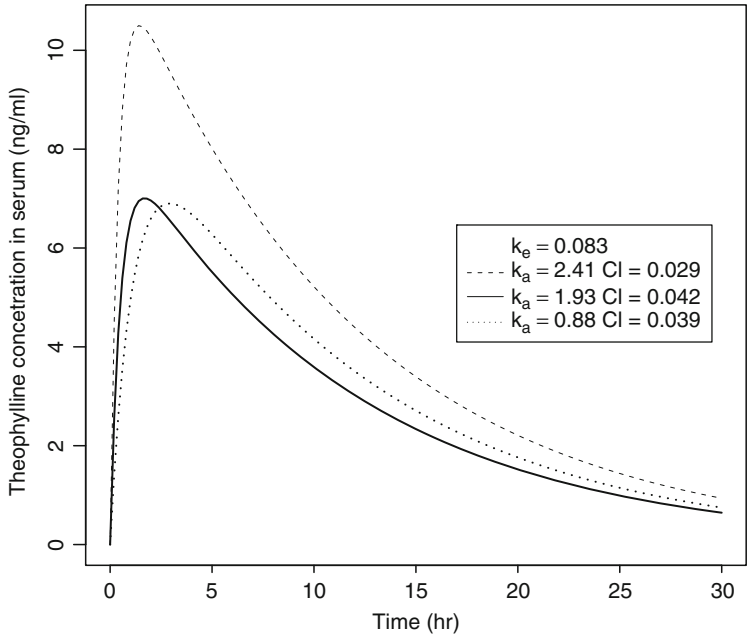

Fig. 2.10 Mixture model for concentration–time curves of 12 subjects receiving an orally administered dose of theophylline

The next subpopulation forms another 48.75% and has a larger absorption constant and higher clearance. Finally, the last subpopulation forms 16.7% of all individuals and is characterized by a large absorption constant $\hat{k}_a = 2.409$ and a small clearance estimated to be 0.028. The corresponding concentration–time curves of the three subpopulations are shown in Fig. 2.10. This model still needs to estimate a large number of parameters, that is, two mixture proportions, one elimination constant, three absorption constants, and three clearance parameters. Thus, a total of nine parameters were estimated for these data. To achieve more data reduction, a normal distribution for the random effects might be considered. In the simplest

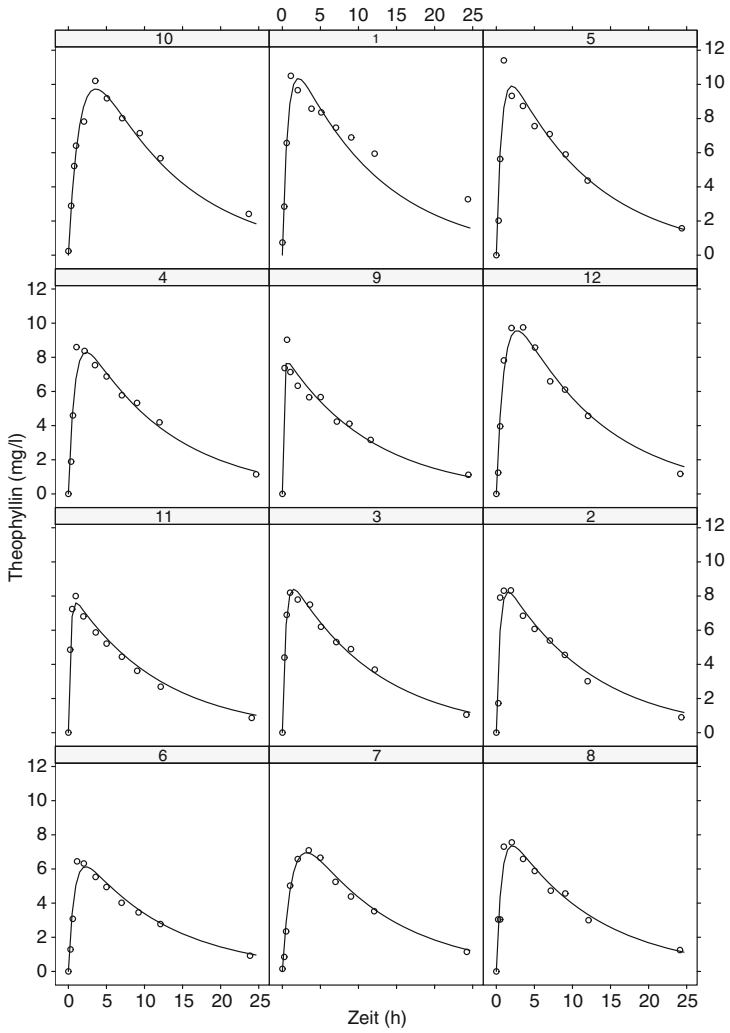

Fig. 2.11 Nonlinear mixed effects model for concentration time curves of 12 subjects receiving an orally administered dose of theophylline

case, e.g., only variability between individuals for the absorption constant would be assumed. This model has fewer parameters, but makes the additional assumption that the variability of the absorption constant between individuals may be described by a normal distribution. However, a satisfactory fit may be achieved if variability between individuals is assumed for the absorption constant and for the clearance. The corresponding estimated concentration–time curves are shown in Fig. 2.11. This model seems to provide an acceptable fit requiring only five parameters.

2.4 A Note of Caution

Finally, at the end of this chapter a note of caution seems in order. Following Box (1979):

> All models are wrong, but some models are useful.

Models are by their very nature only approximations to a complex and complicated reality and thus they are of course literally false. On the other hand, models are the only instruments we have for understanding complex phenomena. The usefulness of the models presented here is hopefully given by the fact that they allow us to identify heterogeneity of individuals and to include known covariates in order to explain some of the diversity between them.

http://www.springer.com/978-3-540-68650-7

Medical Applications of Finite Mixture Models
Schlattmann, P.
2009, X, 246 p. 74 illus., Hardcover
ISBN: 978-3-540-68650-7