

Mixed effects models for longitudinal data

Peter Schlattmann

Charite Universitätsmedizin Berlin

Dept. of Biostatistics and Clinical Epidemiology

University of Southern Denmark, Odense

Dept. of Statistics

Overview

- Linear random effects models for longitudinal data
- Nonlinear random effects models for longitudinal data
- Strategies for model selection based on the likelihood function

Longitudinal data

Orthodontic growth data, Example from Pinheiro and Bates (2000).

Investigators at the University of North Carolina Dental School followed the growth of 27 children (16 males, 11 females) from age 8 until age 14. Every two years they measured the distance between the pituitary and the pterygomaxillary fissure, two points that are easily identified on x-ray exposures of the side of the head.

```
library(nlme)  
plot(Orthodont)
```

The data

Orthodont

Grouped Data: distance ~ age | Subject

	distance	age	Subject	Sex
1	26.0	8	M01	Male
2	25.0	10	M01	Male
3	29.0	12	M01	Male
4	31.0	14	M01	Male
.....				
65	21.0	8	F01	Female
66	20.0	10	F01	Female
67	21.5	12	F01	Female
68	23.0	14	F01	Female

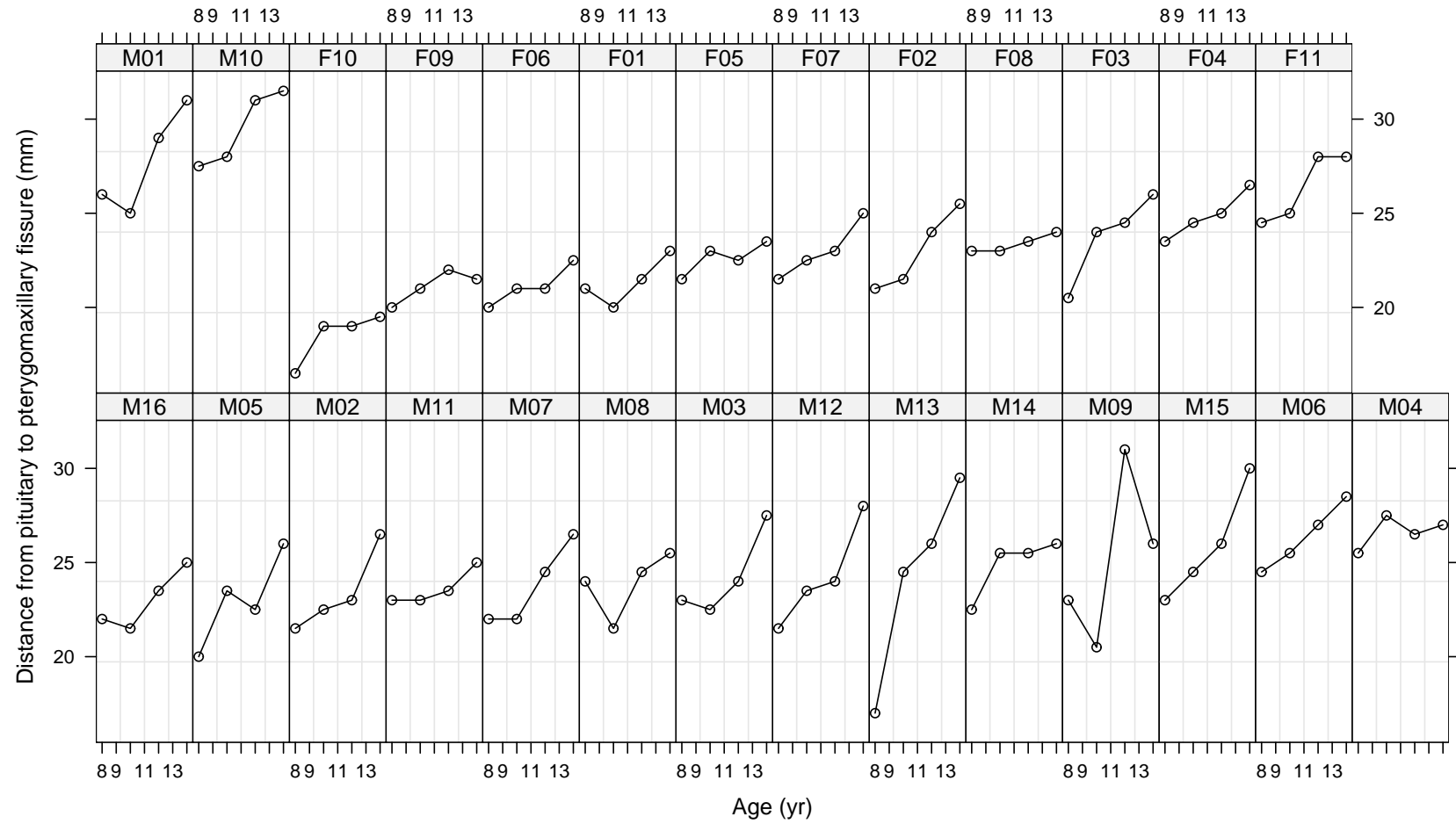


Figure 1: Orthodontic growth data

Questions

- Is there an age effect on growth?
- Is there a gender difference?
- Is growth different in both sexes (Is there an interaction)?
- Is a fixed effects model appropriate?

The simple linear regression model

We have M subjects and n_i measurements per subject

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \epsilon_{ij}, \quad i = 1, \dots, M, \quad j = 1, \dots, n_i$$
$$\epsilon_{ij} \sim N(0, \sigma^2)$$

R call:

```
m0<-glm(distance~Sex+age,data=Orthodont)
```

Result

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	17.70671	1.11221	15.920	< 2e-16	***
SexFemale	-2.32102	0.44489	-5.217	9.20e-07	***
age	0.66019	0.09776	6.753	8.25e-10	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 5.160679)

Null deviance: 917.69 on 107 degrees of freedom
Residual deviance: 541.87 on 105 degrees of freedom
AIC: 488.68

Comment

Bad fit!

Not surprising: we ignored correlation within subjects and variability between subjects

Individual models

Idea: Compute a regression model for each subject!

Restrict analysis to girls (Individual gender differences are nonsense)

R call:

```
Orthof<-Orthodont[Orthodont$Sex=="Female",]  
m1<-lmList(distance~age,data=Orthof)  
postscript("girlsci.eps")  
plot(intervals(m1))  
dev.off()
```

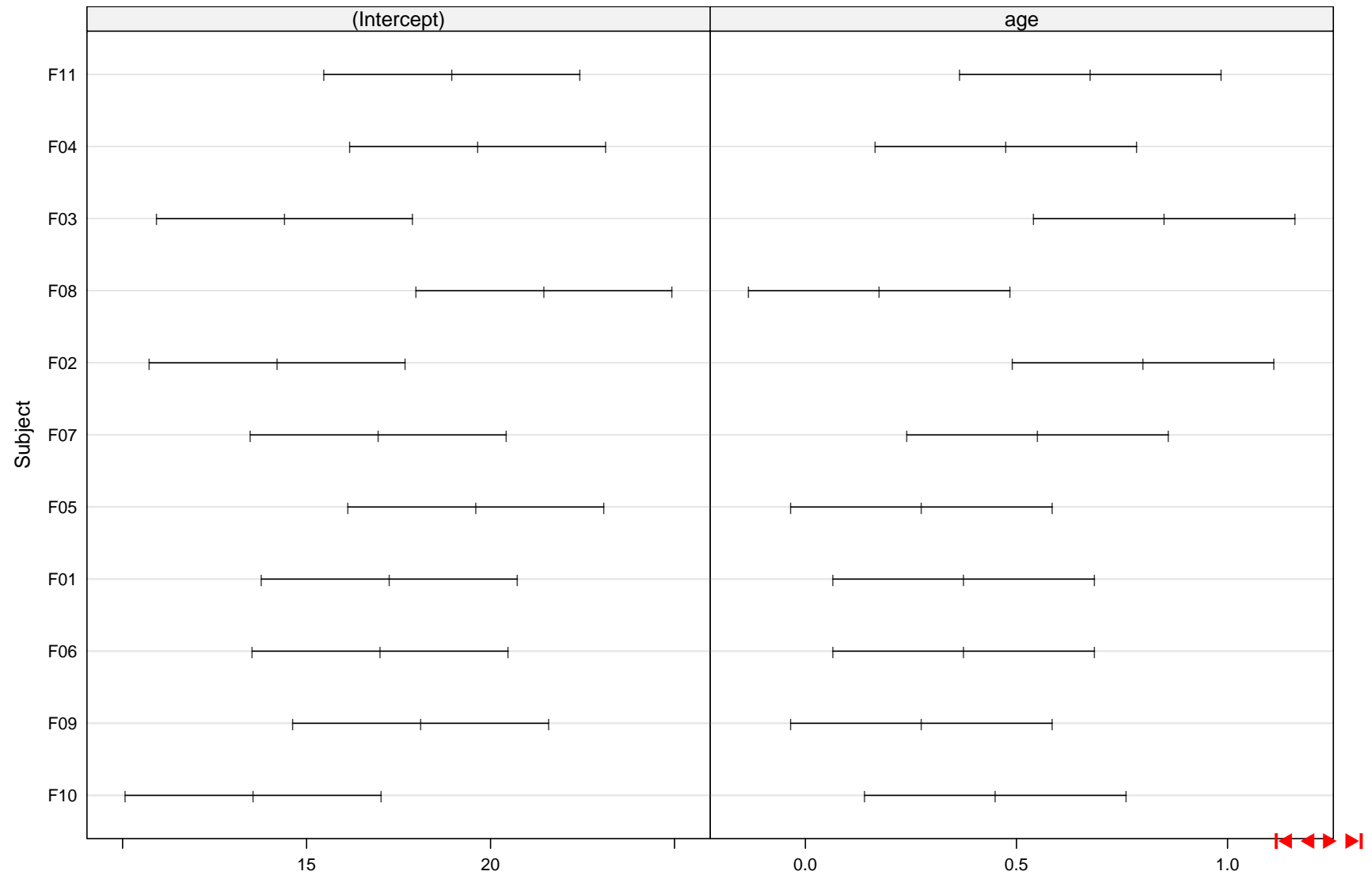


Figure 2: Comparison of 95% confidence intervals

Interpretation

What does the plot show:

We have balanced data, confidence intervals are of similar widths!

What does the plot not show:

Substantial variability between subjects for the intercepts as expected by the plot of the data!

Why is that?

Reason

Age is sampled between 8 and 14 years. The regression models force extrapolation to age 0! The intercept represents a distance at age 0. This induces a high correlation between the estimates of intercepts and slopes.

Alternative: Center the data at age 11!

R call:

```
m2<-lmList(distance~I(age-11),data=Orthof)
postscript("girlsci1.eps")
plot(intervals(m2))
dev.off()
```

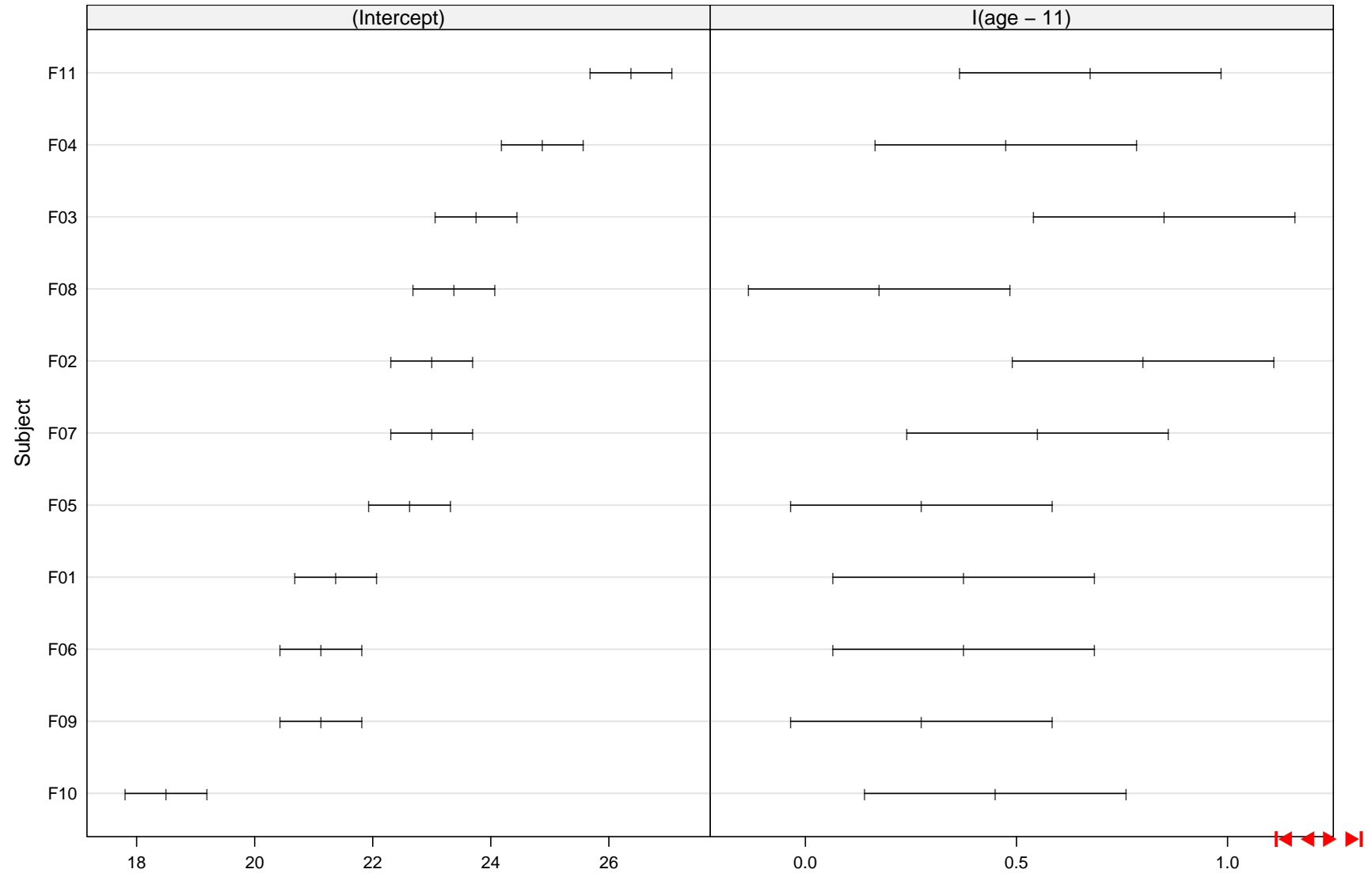


Figure 3: Comparison of 95% confidence intervals

The mixed effects model

We have M subjects and n_i measurements per subject, random effects for intercepts:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + b_i + \epsilon_{ij}, \quad i = 1, \dots, M, \quad j = 1, \dots, n_i$$
$$b_i \sim N(0, \sigma_b^2) \quad \epsilon_{ij} \sim N(0, \sigma^2)$$

in Matrix notation:

$$\mathbf{y}_i = X_i \boldsymbol{\beta} + Z_i \mathbf{b}_i + \boldsymbol{\epsilon}_i$$
$$\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \Sigma) \quad \mathbf{b}_i \sim N(\mathbf{0}, \Psi)$$

Maximum Likelihood Estimation

Hierarchical model:

1. $\mathbf{y}_i \sim N(X_i\boldsymbol{\beta}, \Sigma)$
2. $\mathbf{b}_i \sim N(\mathbf{0}, \boldsymbol{\Psi})$

Taking the marginal density over \mathbf{b}_i leads to the likelihood:

$$L(\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \sigma^2 | \mathbf{y}) = \prod_{i=1}^M \int p(\mathbf{y} | \mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \sigma^2) p(\mathbf{b}) d(\mathbf{b})$$

The Likelihood

$$p(\mathbf{y}|\mathbf{b}, \boldsymbol{\beta}, \sigma^2) = \frac{\exp(-\|\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i\|^2/2\sigma^2)}{(2\pi\sigma^2)^{M/2}}$$
$$p(\mathbf{b}_i) = \frac{\exp(-\mathbf{b}_i^T \boldsymbol{\Psi}^{-1} \mathbf{b}_i/2)}{(2\pi)^{q^2} \sqrt{|\boldsymbol{\Psi}|}}$$

This gives the likelihood:

$$L(\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \sigma^2|\mathbf{y}) \propto \sigma^M |\boldsymbol{\Psi}|^{-M/2} \prod_{i=1}^M \int \exp(-f_i/2) d\mathbf{b}_i$$

$$f_i = \sigma^{-2}((\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i)^T (\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i) + \mathbf{b}_i^T \boldsymbol{\Psi}^{-1} \mathbf{b}_i)$$

Simplification

$$L(\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \sigma^2 | \mathbf{y}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(\frac{-\sum_{i=1}^M \|\tilde{\mathbf{y}}_i - \tilde{\mathbf{X}}_i\boldsymbol{\beta} - \tilde{\mathbf{Z}}_i\hat{\mathbf{b}}_i\|^2}{2\sigma^2}\right) \prod_{i=1}^M \frac{\text{abs}(|\Delta|)}{\sqrt{\tilde{\mathbf{Z}}_i^T \tilde{\mathbf{Z}}_i}}$$

$$\hat{\mathbf{b}}_i = (\tilde{\mathbf{Z}}_i^T \tilde{\mathbf{Z}}_i)^{-1} \tilde{\mathbf{Z}}_i^T (\tilde{\mathbf{y}}_i - \tilde{\mathbf{X}}_i\boldsymbol{\beta})$$

$$\tilde{\mathbf{y}}_i = \begin{bmatrix} \mathbf{y}_i \\ 0 \end{bmatrix} \quad \tilde{\mathbf{X}}_i = \begin{bmatrix} \mathbf{X}_i \\ 0 \end{bmatrix}, \quad \tilde{\mathbf{Z}}_i = \begin{bmatrix} \mathbf{Z}_i \\ \Delta \end{bmatrix}, \quad \Delta = \frac{\boldsymbol{\Psi}^{-1}}{1/\sigma^2}$$

Algorithms

- Direct optimization of (1)
- EM-Algorithm
- Lindstrom-Bates Algorithm
- ...

Example revisited

Fixed effects model with interaction:

R call:

```
m00<-lm(distance~Sex*I(age-11),data=Orthodont)
```

Random effects model over intercepts

R call:

```
m1<-lme(distance~Sex*I(age-11),data=Orthodont,random=~1|Subject)
```

Result: Fixed effects model

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	24.9688	0.2821	88.504	< 2e-16	***
SexFemale	-2.3210	0.4420	-5.251	8.05e-07	***
I(age - 11)	0.7844	0.1262	6.217	1.07e-08	***
SexFemale:I(age - 11)	-0.3048	0.1977	-1.542	0.126	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.257 on 104 degrees of freedom

Multiple R-Squared: 0.4227, Adjusted R-squared: 0.4061

F-statistic: 25.39 on 3 and 104 DF, p-value: 2.108e-12



Result: Random effects model

```
Fixed effects: distance ~ Sex * I(age - 11)
              Value Std.Error DF  t-value p-value
(Intercept)   24.968750 0.4860008 79 51.37595 0.0000
SexFemale     -2.321023 0.7614169 25 -3.04829 0.0054
I(age - 11)    0.784375 0.0775011 79 10.12082 0.0000
SexFemale:I(age - 11) -0.304830 0.1214209 79 -2.51052 0.0141
```

Estimated fixed effects are the same, but different standard errors!

Result: Random effects model

Linear mixed-effects model

Data: Orthodont

AIC	BIC	logLik
445.7572	461.6236	-216.8786

Random effects:

Formula: ~1 | Subject

(Intercept) Residual

StdDev: 1.816214 1.386382

Model update and comparison

New model: Random effects for age

```
m22<-update(m1,random=~I(age-11)|Subject)
```

```
anova(m22,m1,m00)
```

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
m22	1	8	448.5817	469.7368	-216.2908			
m1	2	6	445.7572	461.6236	-216.8786	1 vs 2	1.17559	0.5556
m00	3	5	493.5591	506.7811	-241.7796	2 vs 3	49.80187	<.0001

Random effects for age do not improve the model significantly. The model with random intercepts gives the best fit!

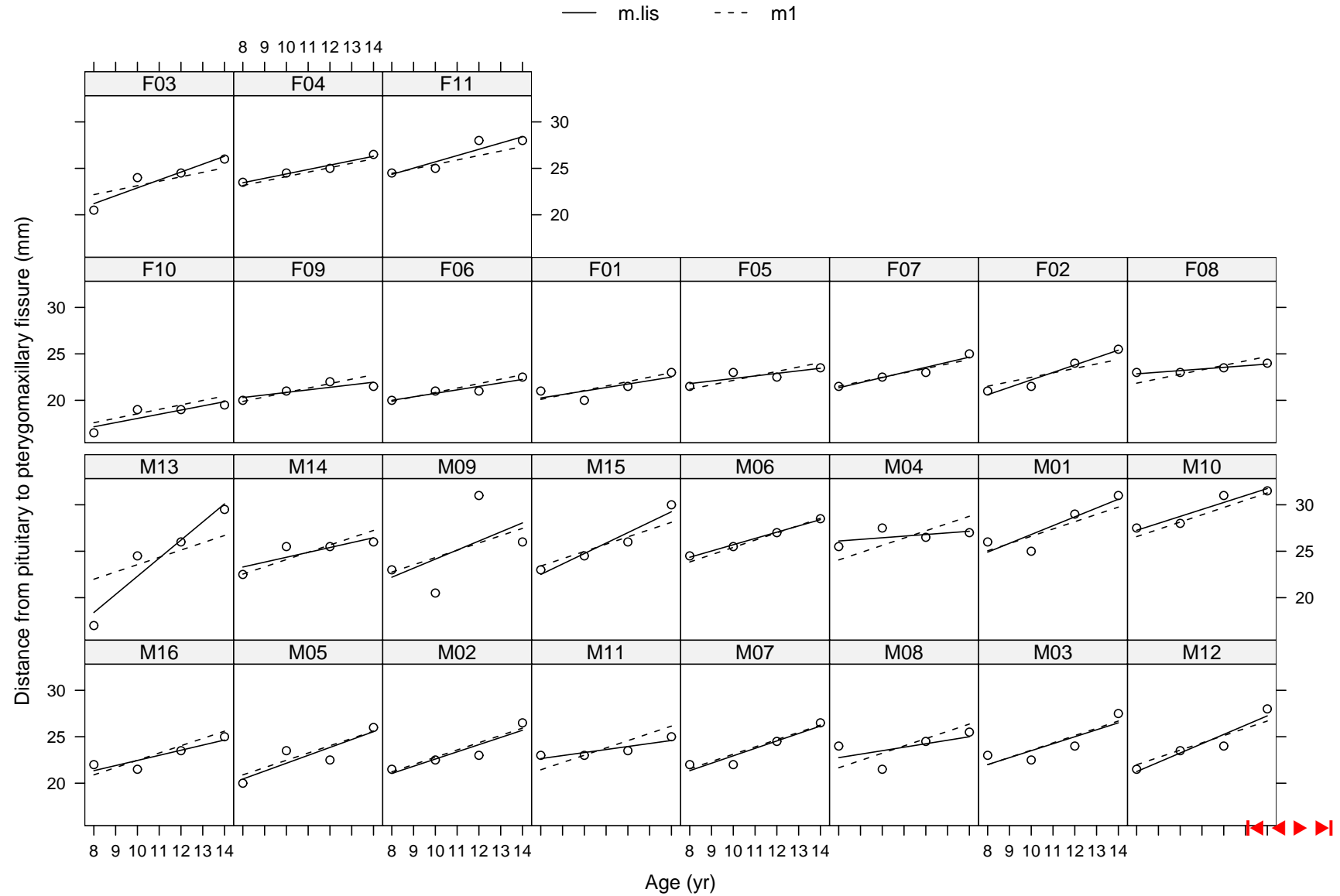


Figure 4: Observed and predicted values

Prediction

```
newOrth<-data.frame(Subject=rep(c("M11","F03"),c(3,3)),  
Sex=rep(c("Male","Female"),c(3,3)), age=rep(16:18,2))
```

```
> newOrth  
  Subject    Sex age  
1     M11  Male  16  
2     M11  Male  17  
3     M11  Male  18  
4     F03 Female  16  
5     F03 Female  17  
6     F03 Female  18
```

Prediction

```
predict(m1,newdata=newOrth,level=0:1)
  Subject predict.fixed predict.Subject
1     M11      28.89062      27.71773
2     M11      29.67500      28.50211
3     M11      30.45937      29.28648
4     F03      25.04545      26.00757
5     F03      25.52500      26.48712
6     F03      26.00455      26.96667
```

For comparison: GEE model I

```
library(gee)
m3<-gee(distance~Sex*I(age-11),id=Subject,family=gaussian(),
corstr = "independence",data=Orthodont)
```

Coefficients:

	Estimate	Naive S.E.	Naive z	Robust S.E.	Robust z
(Intercept)	24.9687500	0.2821186	88.504441	0.44270067	56.400976
SexFemale	-2.3210227	0.4419949	-5.251243	0.74977059	-3.095644
I(age - 11)	0.7843750	0.1261673	6.216945	0.09834755	7.975542
SexFemale:I(age - 11)	-0.3048295	0.1976661	-1.542143	0.11686730	-2.608339

Estimated Scale Parameter: 5.099523



Number of Iterations: 2

Working Correlation

	[,1]	[,2]	[,3]	[,4]
[1,]	1	0	0	0
[2,]	0	1	0	0
[3,]	0	0	1	0
[4,]	0	0	0	1

For comparison: GEE model II

```
gee(formula = distance ~ Sex * I(age - 11), id = Subject, data = Orthodont,  
    family = gaussian(), corstr = "AR-M")
```

Coefficients:

	Estimate	Naive S.E.	Naive z	Robust S.E.	Robust z
(Intercept)	25.0586329	0.4306454	58.188552	0.4384849	57.148217
SexFemale	-2.4159341	0.6746917	-3.580797	0.7544154	-3.202392
I(age - 11)	0.7694569	0.1166051	6.598826	0.1049698	7.330272
SexFemale:I(age - 11)	-0.2856922	0.1826851	-1.563850	0.1223803	-2.334462

Estimated Scale Parameter: 5.099523

Number of Iterations: 2

Working Correlation

	[,1]	[,2]	[,3]	[,4]
[1,]	1.0000000	0.6135188	0.3764054	0.2309318
[2,]	0.6135188	1.0000000	0.6135188	0.3764054
[3,]	0.3764054	0.6135188	1.0000000	0.6135188
[4,]	0.2309318	0.3764054	0.6135188	1.0000000

Within subject correlation in RE models

```
m11<-lme(distance~Sex*age,data=Orthodont,random=~1|Subject,  
correlation=corAR1())
```

Random effects:

```
Formula: ~1 | Subject  
          (Intercept) Residual  
StdDev:    1.826331    1.3731
```

Correlation Structure: AR(1)

```
Formula: ~1 | Subject  
Parameter estimate(s):
```

```
          Phi  
-0.03753262
```

Fixed effects: distance ~ Sex * age

	Value	Std.Error	DF	t-value	p-value
(Intercept)	16.325230	0.9669554	79	16.883126	0.0000
SexFemale	1.051002	1.5149280	25	0.693764	0.4942
age	0.785434	0.0759895	79	10.336085	0.0000
SexFemale:age	-0.306189	0.1190527	79	-2.571873	0.0120

anova(m1,m11)

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
m1	1	6	445.7572	461.6236	-216.8786			
m11	2	7	447.7081	466.2188	-216.8541	1 vs 2	0.04913688	0.8246

Part II - Nonlinear mixed effects models

Pharmacokinetics

Pharmacokinetics is the study of the bodily absorption, distribution, metabolism and excretion of drugs.

Pharmacokinetics looks at the concentration time curve of an individual

Usual types of analysis:

- Compartment models as solutions of differential equations (Gibaldi, 1984)
- Noncompartmental analysis (Cawello, 2000)

Compartment models

Classical models describe the body as a set of interconnected compartments, without a priori anatomical basis. The number of compartments used is the minimum required to fit the pharmacokinetic data available.

- Most simple model - one compartment model
- Transfer between compartments is proportional to the change of concentration (First order kinetics)

Oral dose

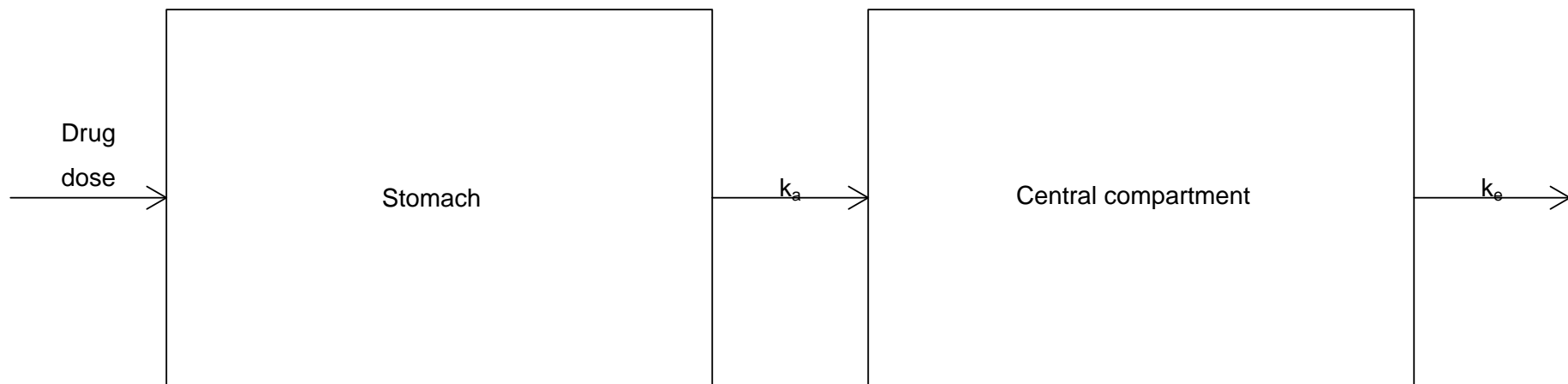


Figure 5: One compartment model with oral dose

Nonlinear model for first order pharmacokinetics

The solution of the corresponding differential equations is given by the nonlinear model:

$$c(t) = \frac{Dk_a}{V(k_a - k_e)}(e^{-k_e t} - e^{-k_a t})$$

$$c(t) = \frac{Dk_a k_e}{Cl(k_a - k_e)}(e^{-k_e t} - e^{-k_a t})$$

Here D is the dose given, V is the volume of distribution, k_a denotes the absorption constant, k_e is the elimination constant and Cl denotes the clearance: The volume of blood completely cleared of a compound per unit time. ($Cl = V * k_e$)

Phase I Clinical Studies

Phase I Studies (First-in-Man):

- Healthy volunteers
- Data rich: Few individuals in a strictly controlled setting
- Frequently analyzed using noncompartmental techniques
- Often ignored: Variability between subjects

Example: Theophylline pharmacokinetics

Theophylline is an oral drug for continuous therapy for patients suffering from Asthma

Phase I Study (Boeckmann et al., 1994):

- 12 healthy volunteers
- Oral dose of Theophylline
- Dose depends on body weight of the individual subject
- Balanced design: 11 concentration measurements over 25 hours

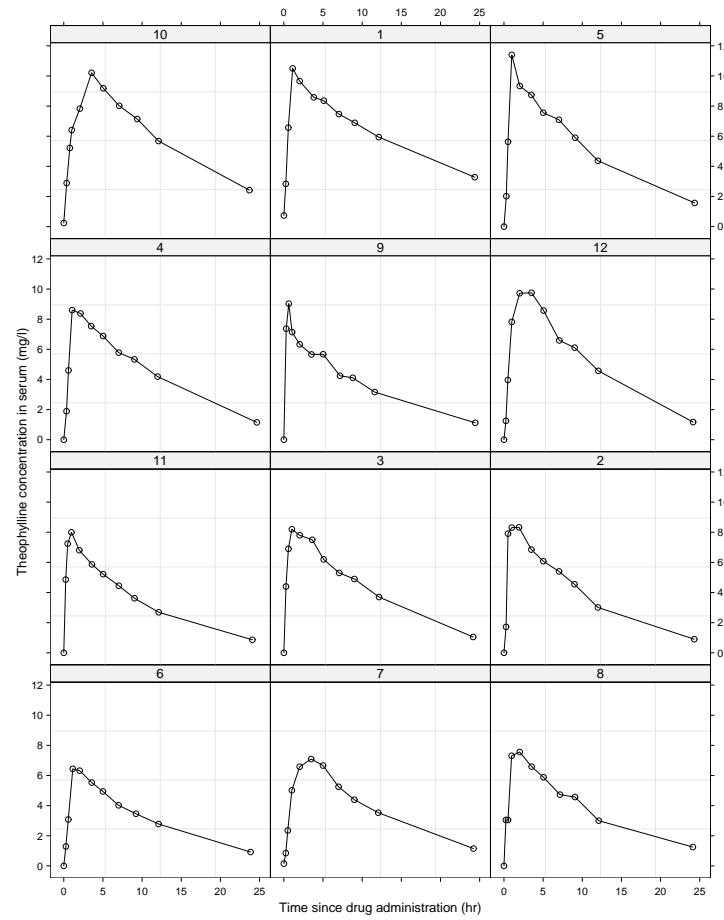


Figure 6: Concentration time curve of Theophylline in 12 healthy volunteers ▶▶▶

Pooled model analysis

$$y_{ij} = \eta(t_{ij}, \lambda) + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma^2)$$

$$\eta(t_{ij}, \lambda) = \frac{D \exp(\beta_1) \exp(\beta_2)}{\exp(\beta_3)(\exp(\beta_2) - \exp(\beta_1))} \left(e^{-\exp(\beta_1)t_j} - e^{-\exp(\beta_2)t_j} \right)$$
$$i = 1, \dots, M, \quad j = 1, \dots, n_i,$$

where $\beta_1 = \log(k_e)$, $\beta_2 = \log(k_a)$, and $\beta_3 = \log(Cl/F)$

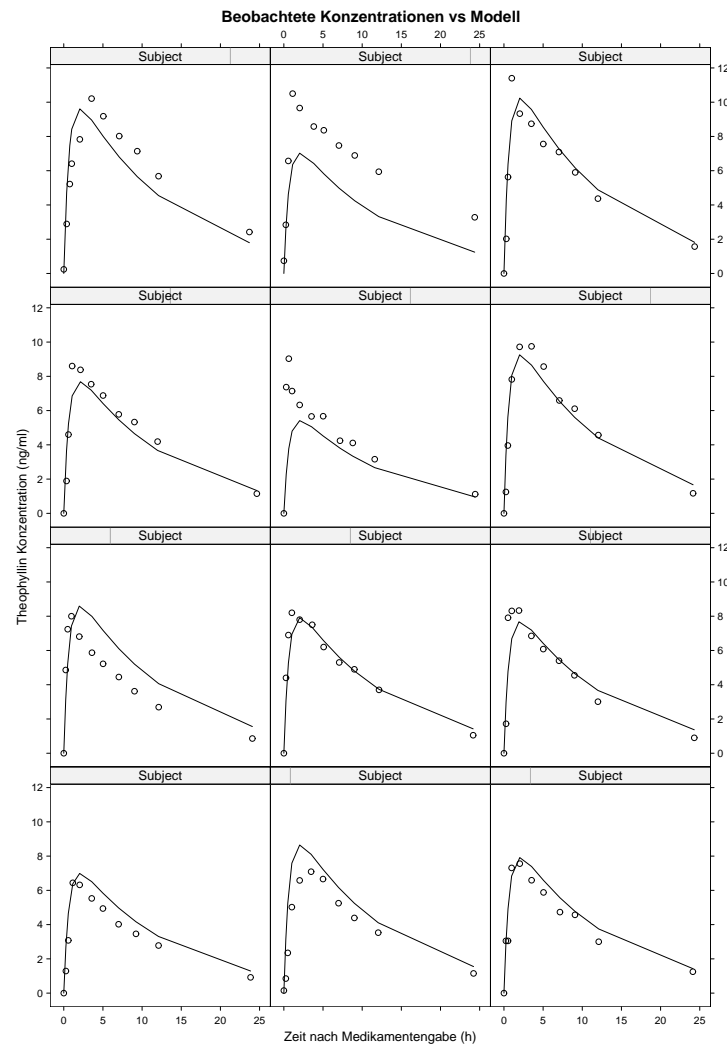


Figure 7: Naive pooled data model



Advantages of simple pooled analysis

- Easy to perform
- Most people in the field are familiar with the model
- Applicable for unbalanced and sparse data

Disadvantages of simple pooled analysis

- No distinction between intra- and interindividual variability
- Unbalanced data may lead to biased results
- Subject specific effects are ignored

Individual model analysis

Idea: Compute a nonlinear model for each subject

$$y_{ij} = \eta(t_{ij}, \lambda_i) + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma^2)$$

$$\eta(t_{ij}, \lambda_i) = \frac{D \exp(\beta_{1i}) \exp(\beta_{2i})}{\exp(\beta_{3i}) (\exp(\beta_{2i}) - \exp(\beta_{1i}))} \left(e^{-\exp(\beta_{1i})t_j} - e^{-\exp(\beta_{2i})t} \right)$$

$i = 1, \dots, M, \quad j = 1, \dots, n_i,$

where $\beta_{1i} = \log(k_{ei})$, $\beta_{2i} = \log(k_{ai})$, and $\beta_{3i} = \log((Cl/F)_i)$

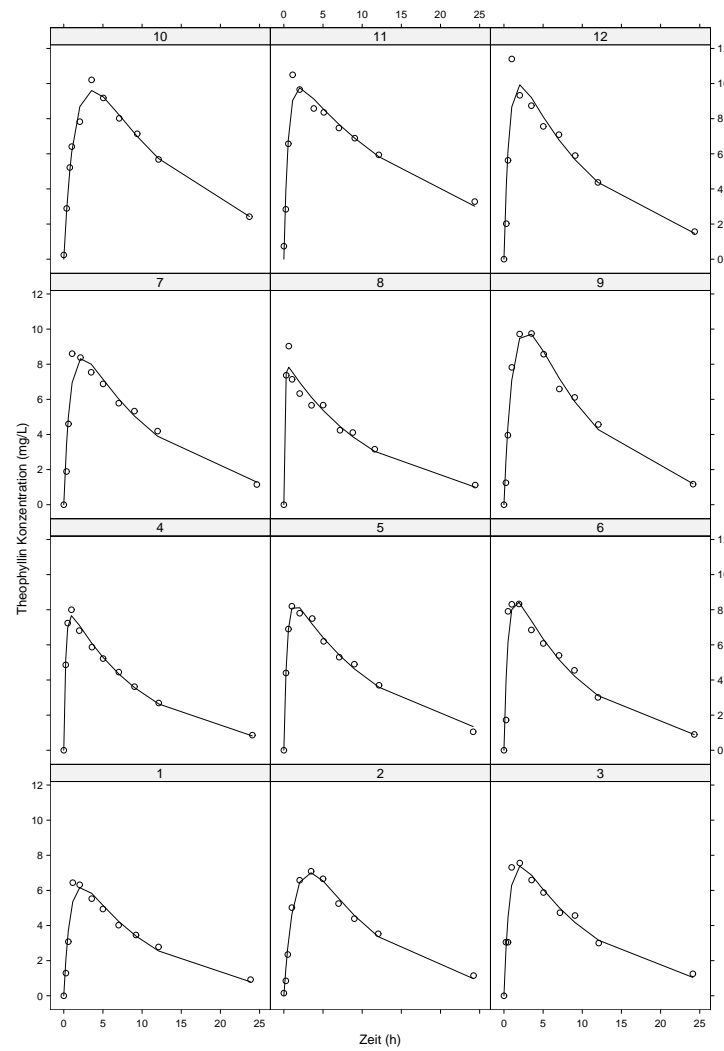


Figure 8: Individual nonlinear models



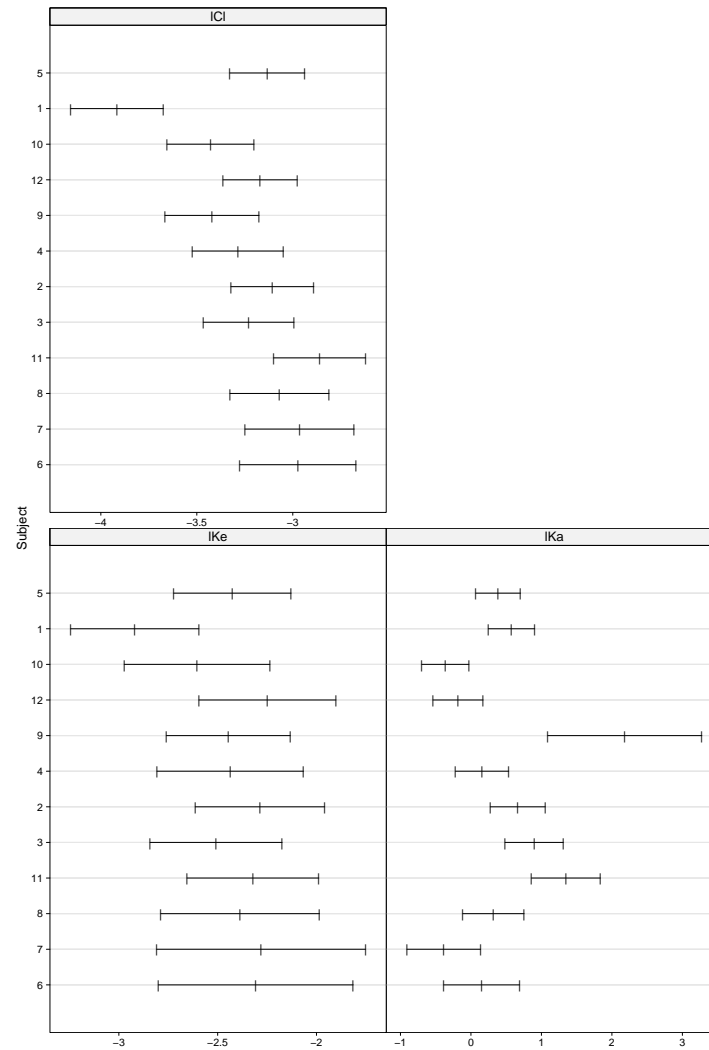


Figure 9: Estimators and confidence intervals of first order pharmacokinetic model

Problems of individual models

- Large number of parameters needs to be estimated
- For sparse data hardly an option
- Only little data reduction
- No quantification of interindividual variability

Population Pharmacokinetics

This approach aims to model the relationship between physiology (both normal and disease altered) and pharmacokinetics while taking into account the interindividual variability in these relationships

Hierarchical models for population pharmacokinetic models

- Intraindividual variability:

$$y_{ij} = \eta(t_{ij}, \lambda_i) + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0, \sigma^2) \quad i = 1, \dots, M \quad j = 1, \dots, n_i$$

- Interindividual variability:

$$\lambda_i \sim P(\mu, \Psi)$$

The distribution P of the random effects is often assumed to be normal (Sheiner, 1974, Lindstrom and Bates, 1990; Davidian and Giltinan, 1995, Lunn, 2002)

Example

Random effects for the elimination constant:

$$y_{ij} = \eta(t_{ij}, \lambda_i, b_{1i}) + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma^2), \quad b_i \sim N(0, \sigma_b^2)$$

$$\eta(t_{ij}, \lambda_i) = \frac{D \exp(\beta_{1i} + b_{1i}) \exp(\beta_{2i})}{\exp(\beta_{3i})(\exp(\beta_{2i}) - \exp(\beta_{1i} + b_{1i}))} \left(e^{-\exp(\beta_{1i} + b_{1i})t_j} - e^{-\exp(\beta_{2i})t} \right)$$

$$i = 1, \dots, M, \quad j = 1, \dots, n_i,$$

where $\beta_{1i} = \log(k_{ei})$, $\beta_{2i} = \log(k_{ai})$, and $\beta_{3i} = \log((Cl/F)_i)$



Maximum likelihood estimation

The marginal density of \mathbf{y} is given by:

$$p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\Psi}, \sigma^2) = \int p(\mathbf{y}|\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \sigma^2)p(\mathbf{b})d(\mathbf{b})$$

As before the likelihood is given by:

$$L(\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \sigma^2|\mathbf{y}) = \prod_{i=1}^M \int p(\mathbf{y}|\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \sigma^2)p(\mathbf{b})d(\mathbf{b})$$

Likelihood

$$p(\mathbf{y}|\mathbf{b}, \boldsymbol{\beta}, \sigma^2) = \frac{\exp(-\|\mathbf{y}_i - \boldsymbol{\eta}(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i)\|^2/2\sigma^2)}{(2\pi\sigma^2)^M/2}$$

$$p(\mathbf{b}_i) = \frac{\exp(-\mathbf{b}_i^T \boldsymbol{\Psi}^{-1} \mathbf{b}_i)/2}{(2\pi)^{q^2} \sqrt{|\boldsymbol{\Psi}|}}$$

$$L(\mathbf{b}, \boldsymbol{\beta}, \boldsymbol{\Psi}, \sigma^2|\mathbf{y}) \propto \sigma^M |\boldsymbol{\Psi}|^{-M/2} \prod_{i=1}^M \int \exp(-f_i/2) d\mathbf{b}_i$$

$$f_i = \sigma^{-2}(\mathbf{y}_i - \boldsymbol{\eta}(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i))^T (\mathbf{y}_i - \boldsymbol{\eta}(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i)) + \mathbf{b}_i^T \boldsymbol{\Psi}^{-1} \mathbf{b}_i$$

Approximations of the likelihood

The random effects are a nonlinear part of the integrand. Usually there is no closed form solution. The following approximations are frequently used:

- Taylor series development at $e_{ij} = 0$ $\mathbf{b}_i = 0$ (First order model, Sheiner and Beal, 1982)
- Taylor series development at the conditional modes of $\hat{\mathbf{b}}_i$ (Lindstrom and Bates, 1990)
- Gaussian quadrature (Davidian and Giltinan, 1992)
- Monte Carlo integration (Pinheiro and Bates, 1995)

- Alternative: Bayesian approaches (Wakefield, 1996; Lunn 2002)

Choice of fixed and random effects terms

Idea: Use individual estimates β_i and corresponding confidence intervals

Interpretation:

- Small variability of lK_e : fixed effect

- High variability of lK_a and of clearance lCl : Random Effects

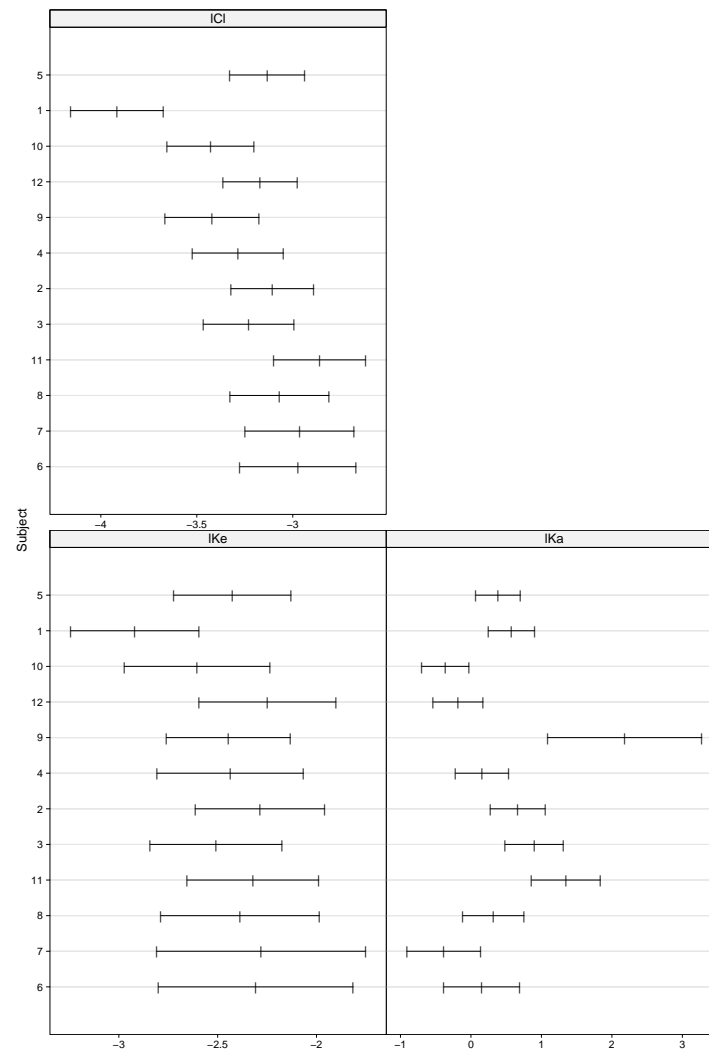


Figure 10: Estimators and confidence intervals



Result: Nonlinear mixed effects model I

```
Fixed: list(lKe ~ 1, lKa ~ 1, lCl ~ 1)
          lKe          lKa          lCl
-2.4547038  0.4657339 -3.2272226
```

Random effects:

```
Formula: list(lKa ~ 1, lCl ~ 1)
Level: Subject
Structure: Diagonal
          lKa          lCl  Residual
StdDev: 0.6435777 0.1669281 0.7092541
```

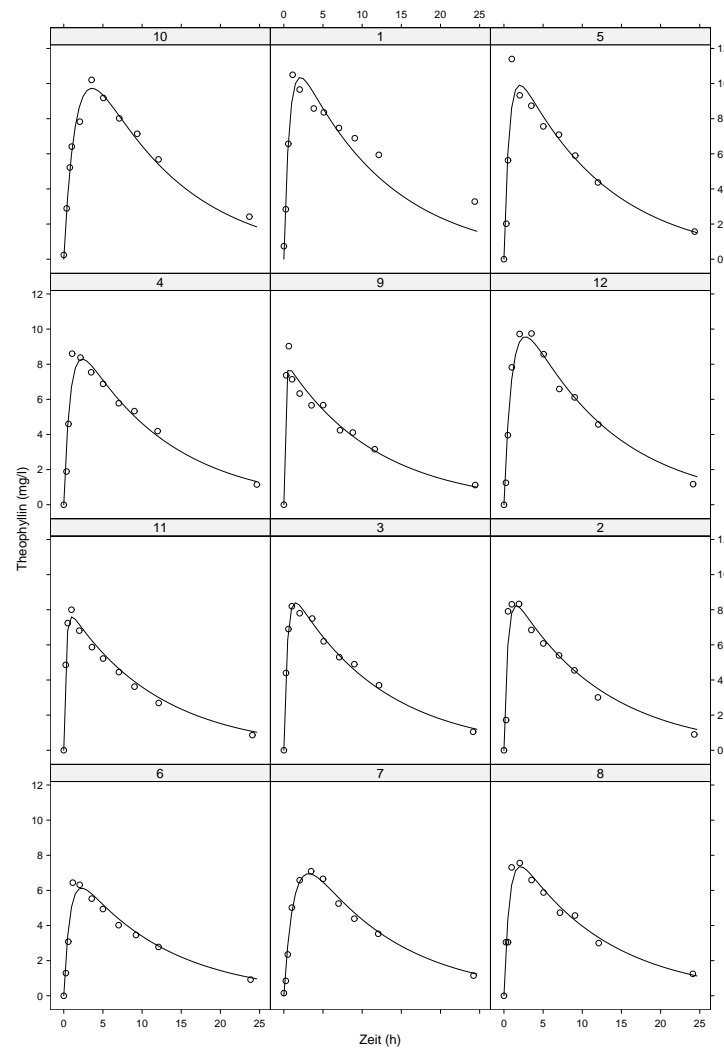


Figure 11: Nonlinear mixed effects model



Result: Nonlinear mixed effects model II

Log-likelihood: -216.5335

Fixed: list(lKe ~ 1, lKa ~ 1, lCl ~ 1)

lKe	lKa	lCl
-2.5441213	0.4370874	-3.2753672

Random effects:

Formula: lCl ~ 1 | Subject

lCl Residual

StdDev: 0.1721872 1.130623

Model selection via AIC and BIC

Likelihood with penalty term: Akaike's and Bayesian information criterion:

Model	df	AIC	BIC	logLik
Fixed	4	479.2190	490.7502	-235.6095
+1Cl	5	443.0669	457.4809	-216.5335
+1Cl+1Ka	6	366.0429	383.3397	-177.0214

Choice of distribution for the random effects

Instead of a normal distribution for the random effects a finite mixture model may be used.

Finite mixture model

Conditional distribution of Y

$$Y_i \sim p_1 f(y_i, \lambda_1) + \cdots + p_k f(y_i, \lambda_k) \quad k \text{ components}$$

$$f(y_i, \lambda_j) = \prod_l^{n_i} N(y_{il}, \eta(t_{il}, \lambda_j), \sigma) \quad n_i \text{ observations}$$

Assuming conditional independence for each individual the density f is given by the product of the normal densities of this individual

Each mixture component is described by a separate nonlinear regression model with parameter vector λ_j .

Finite mixture models

Discrete finite random parameter P

$$P = \begin{bmatrix} \lambda_1 & \dots & \lambda_k \\ p_1 & \dots & p_k \end{bmatrix}$$

with $\lambda_j = (\beta_{j1}, \dots, \beta_{jm})^T$ m parameters

and $j = 1, \dots, k$ components

Coefficients λ_j , mixing weights p_j and the number of components k must be estimated from the data, e.g. using maximum likelihood using the EM-algorithm.

This model has the advantage that the random effects distribution does not need to be approximated!

Conclusion

Random effects models provide a flexible tool for modeling.

Available in standard software

- R/S-plus (nlme)
- SAS (proc mixed, proc nlmixed)
- Stata

Crucial assumption: The type of RE-distribution (must it be always the normal distribution?)

Final slide

Thank you for your attention!

