Models for longitudinal data
Analysis of repeated measurements, 28th November 2014

Julie Lyng Forman & Lene Theil Skovgaard
Department of Biostatistics, University of Copenhagen

Topics for today

More models for longitudinal data
- Models for the mean
- Models for the covariance
- Random regression

Read: Fitzmaurice et al. (2011) chapters 6, 7, and 8.

Baseline adjustment
- in randomized clinical trials
- in observational studies

Read: Fitzmaurice et al. (2011), section 5.6.
Papers by Vicker's & Altman and Liu et al (see relevant section).

Outline

Longitudinal analysis
- Models for the mean
- Covariance pattern models
- Random regression
- Baseline adjustments

Typical set-up for longitudinal measurements

Two or more groups of subjects
- Often receiving different treatments
- Some with randomization at baseline.

Longitudinal measurements of the same quantity over time for each subject, typically as a function of
- time (i.e. duration of treatment)
- age
- cumulative dose of drug

Measurements made on the same person are correlated.
- This must be accounted for, since otherwise statistical inference will be biased.
Longitudinal designs

Merits
▶ Are much more powerful in detecting change over time. (data are 'paired' with the subject as its own control)
▶ We may discover that subjects have different time courses (In designs with only cross-sectional data, this may also be the case, but we have no way of knowing!)

Drawbacks
▶ Analysis by traditional ANOVA or GLM-models is infeasible because the independence assumption is violated.
▶ We need a model for the mean and for the covariance.

Repetition: Analysis of response profiles
Comparison of change over \( n \) time points (time) within \( g \) groups (treatment) of subjects.
▶ Similar to two-way ANOVA only with correlated data.
▶ An unstructured covariance is assumed.

Drawbacks:
▶ can only handle balanced designs.
▶ is no good with many groups or time points, due to too many model parameters.
▶ Do not make use of a priori known data patterns, e.g.
  ▶ correlation decreasing with time.
  ▶ Steadily increasing response to treatment.

Today: Parametric models for the mean and the covariance.

Outline
Longitudinal analysis
Models for the mean
Covariance pattern models
Random regression
Baseline adjustments
Models for the mean

Most often changes over time appear **gradually**.

- **We gain power by incorporating this in our models.**

Models the mean as a **continuous function of time**:

- **Linear**
- **Polynomial**
- **Piecewise linear (i.e. spline)**
- **(Cyclic)**
- **(Nonparametric)**

---

**Cardiac: spline model**

Fit at linear spline with break point at $t_2 = 24$ months.

- $\text{ctime} = \text{time} \times \text{continuous version of time}$
- IF time $\leq 24$ THEN $\text{xtime2y} = 0$;
- IF time $> 24$ THEN $\text{xtime2y} = (\text{time} - 24)$;

PROC MIXED DATA=cardiacx METHOD=ML; WHERE time > 0;
CLASS id treatment time;
MODEL logco = ctime treatment treatment*ctime xtime2y xtime2y*treatment / SOLUTION DDFM=SATTERTHWAITE OUTPM=fit_spline;
REPEATED time / subject=id type=un;
RUN;

**Note:** Estimated mean response profiles are stored in fit_spline.

---

**Cardiac: estimated response profiles**

Hypothesis: **No or little further improvement beyond 2 years**

---

**Cardiac: Estimates**

**Fit Statistics**

- $-2 \log \text{Likelihood} = 30.9$ <-------- Used later
- AIC (smaller is better) = 23.1
- AICC (smaller is better) = 44.4
- BIC (smaller is better) = 48.6

**Solution for Fixed Effects**

| Effect          | treatment | Estimate | StdError | DF  | t Value | Pr > |t| |
|-----------------|-----------|----------|----------|-----|---------|------|---|
| Intercept       |           | 1.4903   | 0.09132  | 19.4| 16.32   | <.0001 |
| ctime           | treatment | -0.01627 | 0.00377  | 20  | -4.31   | 0.0003 |
| treatment kombi|           | -0.2984  | 0.1167   | 18.5| -2.56   | 0.0195 |
| treatment mono |           | 0        |          |     |         |      |   |
| ctime*treatment| kombi     | 0.02715  | 0.00478  | 19.1| 5.67    | <.0001 |
| ctime*treatment| mono      | 0        |          |     |         |      |   |
| xtime2y         | treatment | 0.03132  | 0.000708 | 21  | 4.42    | 0.0002 |
| xtime2y*treatment| kombi| -0.03952 | 0.00904  | 19.9| -4.37   | 0.0003 |
| xtime2y*treatment| mono| 0        |          |     |         |      |   |

The slope of the mono-group is $-0.016 (0.004)$ before 2 years. It is $0.027 (0.005)$ lower than in the komb-group. After two years the slope increases by $+0.031 (0.007)$ in the mono-group, while it decreases a bit for the kombi-group (by $0.0313-0.0395 = -0.008$).
Cardiac: Estimated response profiles
(kombination therapy vs mono therapy.

BUT: Is the spline-model at all plausible?

Conventional or residual likelihood?

When testing parametric submodels for the mean.

- Only the conventional likelihood is valid.

The deviance of the spline model must be compared to:

PROC MIXED DATA=cardiacx METHOD=ML; WHERE time > 0;
CLASS id treatment reftime time;
MODEL logco = treatment reftime reftime*treatment /
  SOLUTION DDFM=SATTERTHWAITE;
REPEATED time / subject=id type=un;
RUN;

But don’t forget: Most hypothesis about the mean can be tested using just the default F-tests (optimal choice).

Comparison of models for the mean

Use the likelihood – the ML-version:

- Better fitting models have large values of likelihood $L$ and therefore small values of deviance: $-2 \log L$
- Compute differences in deviances ($\Delta = -2 \log Q$) and compare to a $\chi^2$-distribution with $df = \Delta$ no. params.

Note: Only nested models can be compared.

- The unrestricted 2-way ANOVA can be used as reference as it contains all other submodels.

Example: Spline vs response profiles.

$$-2 \log Q = 38.3 - 30.9 = 7.4 \sim \chi^2(12 - 6) = \chi^2(6) \Rightarrow P = 0.29$$

Model comparisons

Cardiac example: Simplified models for the mean.

<table>
<thead>
<tr>
<th>Model</th>
<th>$-2 \log L$</th>
<th>no. parms</th>
<th>$-2 \log Q$</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saturated</td>
<td>-38.3</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spline</td>
<td>-30.9</td>
<td>6</td>
<td>7.4</td>
<td>6</td>
<td>0.29</td>
</tr>
<tr>
<td>Quadratic</td>
<td>-24.3</td>
<td>6</td>
<td>14.0</td>
<td>6</td>
<td>0.030</td>
</tr>
<tr>
<td>Linear</td>
<td>-17.7</td>
<td>4</td>
<td>20.6</td>
<td>8</td>
<td>0.008</td>
</tr>
</tbody>
</table>
What about the treatment effect?
Get the deviance of the spline-model without treatment-effects:

```
PROC MIXED DATA=cardiacx METHOD=ML; WHERE time > 0;
CLASS id treatment reftime;
MODEL logco = time xtime2y / SOLUTION DDFM=SATTERTHWAITE;
REPEATED reftime / subject=id type=un;
RUN;
```

Fit Statistics

−2 Log Likelihood −7.4

Likelihood ratio test:

\[-2 \log Q = 30.9 - 7.4 = 23.5 \sim \chi^2(6 - 3) = \chi^2(3) \Rightarrow P < 0.0001\]

Example: Calcium supplements

A total of 112 11-year old girls were randomized to receive either calcium or placebo.

Outcome: BMD = bone mineral density, in \( \frac{g}{cm^2} \)

Follow-up: every 6 months, 5 visits in total including baseline

Does calcium improve the bone gain for adolescent women?
Analysis of response profiles

At first do not assume any specific structure neither for the mean nor for the covariances.

▶ Use the saturated model as reference point.

PROC MIXED DATA=calcium;
CLASS grp girl visit;
MODEL bmd=grp visit grp*visit / DDFM=SATTERTHWAITE;
REPEATED visit / TYPE=UN SUBJECT=girl(grp) R RCORR;
RUN;

The unstructured covariance

Advantages
▶ We make no wrong assumptions about the covariance of our observations.
▶ We gain insight in the actual structure of the covariance.

Drawbacks
▶ We use quite a lot of parameters to describe the covariance structure. Thus our analysis becomes less efficient.
▶ No good with small data sets: The results may be unstable.
▶ It can only be used in case of balanced data, i.e. all subjects have to be measured at identical times.

Models for the covariance

Most often covariance display distinct features.

E.g. decreasing correlation with increasing time span between observations …as in the calcium data!

▶ We gain power by incorporating these features in our model.

Possibilities:
▶ Unstructured covariance
▶ Covariance pattern models
▶ Variance components / random effects
(all available with proc mixed)

Output

Estimated R Correlation Matrix for girl(grp) 101 C

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
<th>Col5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>0.9699</td>
<td>0.9414</td>
<td>0.9250</td>
<td>0.8987</td>
</tr>
<tr>
<td>2</td>
<td>0.9699</td>
<td>1.0000</td>
<td>0.9727</td>
<td>0.9585</td>
<td>0.9399</td>
</tr>
<tr>
<td>3</td>
<td>0.9414</td>
<td>0.9727</td>
<td>1.0000</td>
<td>0.9809</td>
<td>0.9592</td>
</tr>
<tr>
<td>4</td>
<td>0.9250</td>
<td>0.9585</td>
<td>0.9809</td>
<td>1.0000</td>
<td>0.9755</td>
</tr>
<tr>
<td>5</td>
<td>0.8987</td>
<td>0.9399</td>
<td>0.9592</td>
<td>0.9755</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Fit Statistics

-2 Res Log Likelihood -2346.3 <----------used later

Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>grp</td>
<td>1</td>
<td>109</td>
<td>2.55</td>
<td>0.1129</td>
</tr>
<tr>
<td>visit</td>
<td>4</td>
<td>97.1</td>
<td>258.08</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>grp*visit</td>
<td>4</td>
<td>97.1</td>
<td>2.79</td>
<td>0.0303</td>
</tr>
</tbody>
</table>
Stationary covariance patterns

Most available models for are for equidistant observations, assuming both variances and correlations are stationary, i.e.

- The variances are all the same
- Correlation depend only on the time-distance between the observations.

\[
\text{proc mixed}
\]

\[
\begin{array}{|c|c|c|}
\hline
\text{type=} & \text{Cov}(Y_{ij}, Y_{ik}) & \text{no. par} \\
\hline
\text{CS} & \sigma^2 I\{j = k\} & 2 \\
\text{AR(1)} & \sigma^2 \rho^{k-j} & 2 \\
\text{ARMA(1,1)} & \sigma^2 [I\{j = k\} + \gamma \cdot \rho^{k-j-1} I\{j \neq k\}] & 3 \\
\text{TOEP} & \sigma^2 [I\{j = k\} + \rho \cdot I\{j \neq k\}] & n \\
\hline
\end{array}
\]

aka the compound symmetry, autoregressive, autoregressive moving average, and the Toeplitz models.

Heterogeneous covariance patterns

The assumption that the variances are stationary can be dropped in which case we have a heterogeneous model for the variances.

- No restrictions on the variances
- Correlation depend only on the time-distance between the observations.

\[
\text{proc mixed}
\]

\[
\begin{array}{|c|c|c|}
\hline
\text{type=} & \text{Cov}(Y_{ij}, Y_{ik}) & \text{no. par} \\
\hline
\text{CSH} & \sigma_j \sigma_k I\{j = k\} + \rho \cdot I\{j \neq k\} & n + 1 \\
\text{ARH(1)} & \sigma_j \sigma_k \rho^{k-j} & n + 1 \\
\text{TOEPH} & \sigma_j \sigma_k [I\{j = k\} + \rho_{j-k} \cdot I\{j \neq k\}] & 2n - 1 \\
\text{ANTE(1)} & \sigma_j \sigma_k \prod_{l=1}^{k-1} \rho_l & 2n - 1 \\
\hline
\end{array}
\]

aka the heterogeneous compound symmetry, heterogeneous autoregressive, the heterogeneous Toeplitz, and the antedependence covariance structures.

Covariance matrix - AR(1)

\[
\text{Cov}(Y_{ij}, Y_{ik}) = \sigma^2 \cdot \rho^{k-j}
\]

\[
\Sigma = \sigma^2 \cdot \begin{pmatrix}
1 & \rho & \rho^2 & \rho^3 & \rho^4 \\
\rho & 1 & \rho & \rho^2 & \rho^3 \\
\rho^2 & \rho & 1 & \rho & \rho^2 \\
\rho^3 & \rho^2 & \rho & 1 & \rho \\
\rho^4 & \rho^3 & \rho^2 & \rho & 1
\end{pmatrix}
\]

- Constant variance $\sigma^2$.
- Correlation decreasing as a power with the distance between the observations.

Autoregressive covariance structure in SAS

Fit the calcium data with:

```
PROC MIXED DATA=calcium;
CLASS grp girl visit;
MODEL bmd=grp visit grp*visit / DDFM=SATTERTHWAITE;
REPEATED visit / TYPE=AR(1) SUBJECT=girl(grp) R RCORR;
RUN;
```

Note: Similar syntax is valid for other types of covariance patterns (see the tables above).
Output from TYPE=AR(1) structure

Estimated R Correlation Matrix for girl(grp) 101 C

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
<th>Col5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.000</td>
<td>0.9708</td>
<td>0.9425</td>
<td>0.9150</td>
<td>0.8883</td>
</tr>
<tr>
<td>2</td>
<td>0.9708</td>
<td>1.0000</td>
<td>0.9708</td>
<td>0.9425</td>
<td>0.9150</td>
</tr>
<tr>
<td>3</td>
<td>0.9425</td>
<td>0.9708</td>
<td>1.0000</td>
<td>0.9708</td>
<td>0.9425</td>
</tr>
<tr>
<td>4</td>
<td>0.9150</td>
<td>0.9425</td>
<td>0.9708</td>
<td>1.0000</td>
<td>0.9708</td>
</tr>
<tr>
<td>5</td>
<td>0.8883</td>
<td>0.9150</td>
<td>0.9425</td>
<td>0.9708</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>girl(grp)</td>
<td>0.9708</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>0.004412</td>
</tr>
</tbody>
</table>

Fit Statistics

-2 Res Log Likelihood = -2318.6

Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>DF</th>
<th>DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>grp</td>
<td>1</td>
<td>113</td>
<td>2.74</td>
<td>0.1006</td>
</tr>
<tr>
<td>visit</td>
<td>4</td>
<td>382</td>
<td>233.91</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>grp*visit</td>
<td>4</td>
<td>382</td>
<td>2.86</td>
<td>0.0232</td>
</tr>
</tbody>
</table>

Comparison of covariance structures

Use the likelihood – either of REML or ML will do:

- Better fitting models have large values of likelihood $L$ and therefore small values of deviance: $-2 \log L$
- Compute differences in deviances ($\Delta = -2 \log Q$) and compare to a $\chi^2$-distribution with $df = \Delta$ no. params.

Note: Only nested models can be compared.
- We can use the unstructured covariance as reference point since it contains all other models as submodels.

Example: AR vs UN

$-2 \log Q = 2346.3 - 2318.6 = 27.7$
$\sim \chi^2(15 - 2) = \chi^2(13) \Rightarrow P = 0.01$

Better stick to the unstructured covariance.

(Or try ARH(1) since the variances seem to increase with time).

Comparison of covariance structures

<table>
<thead>
<tr>
<th>Model</th>
<th>-2 \log L</th>
<th>par.</th>
<th>-2\log Q</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN</td>
<td>-2346.3</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARMA(1,1)</td>
<td>-2318.6</td>
<td>3</td>
<td>27.7</td>
<td>12</td>
<td>0.006</td>
</tr>
<tr>
<td>AR(1)</td>
<td>-2318.6</td>
<td>2</td>
<td>27.7</td>
<td>13</td>
<td>0.010</td>
</tr>
<tr>
<td>CS</td>
<td>-2188.8</td>
<td>2</td>
<td>129.8</td>
<td>13</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

Predicted mean time profiles

Note: Estimated profiles are almost identical for all choices of covariance structures. In fact, for balanced data, they agree completely (since they are equal to the group*time-averages).
Tests of treatment effect

**BUT**: We cannot make inference from profiles alone

- Confidence intervals and tests depend on the covariance.

<table>
<thead>
<tr>
<th>Covariance structure</th>
<th>Test statistic $\sim$ distribution</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independence</td>
<td>$0.35 \sim F(4,491)$</td>
<td>0.84</td>
</tr>
<tr>
<td>Compound symmetry</td>
<td>$5.30 \sim F(4,382)$</td>
<td>0.0004</td>
</tr>
<tr>
<td>Autoregressive</td>
<td>$2.86 \sim F(4,382)$</td>
<td>0.023</td>
</tr>
<tr>
<td>Unstructured</td>
<td>$2.72 \sim F(4,107)$</td>
<td>0.034</td>
</tr>
</tbody>
</table>

*P-values for the interaction term group*visit.

Modeling strategy

As suggested by *Fitzmaurice et al.* (2011):

1. Put up a plausible (e.g. saturated) model for the mean
2. Fit the data so far ignoring correlation (GLM).
3. Check the residuals for assessing the adequacy of the model for the mean and in order to get an impression of the error covariance.
4. Pick a reasonable model for the covariance (if possible test against the unstructured model).
5. Re-check the model fit.
6. Do the analysis.

**Note**: Lecture 5 on model diagnostics among others.

Outline

- Longitudinal analysis
- Models for the mean
- Covariance pattern models
- Random regression
- Baseline adjustments

Calcium data

The time course looks reasonably linear, but maybe the individual girls have different growth rates . . .
Individual regression

Fit an ordinary linear regression for each girl:

\[ Y_{ij} = A_i + B_i t_j + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, \sigma^2) \]

Are the slopes of the Calcium-group bigger?

Summarizing individual regressions

Analysis of summary statistics

Estimated intercepts and slopes (se):

<table>
<thead>
<tr>
<th>Group</th>
<th>Level at first visit</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>0.8697 (0.0086)</td>
<td>0.0206 (0.0014)</td>
</tr>
<tr>
<td>C</td>
<td>0.8815 (0.0088)</td>
<td>0.0244 (0.0014)</td>
</tr>
<tr>
<td>Difference</td>
<td>0.0118 (0.0123)</td>
<td>0.0039 (0.0019)</td>
</tr>
</tbody>
</table>

P-value: 0.34 0.050

\* two-sample t-test.

Random regression

We let each girl have her own level \( A_i \) and her own slope \( B_i \)

We assume these individual 'parameters' \( A_i \) and \( B_i \) follow a bivariate normal distribution in the population

\[
\begin{pmatrix} A_i \\ B_i \end{pmatrix} \sim N_2\left( \begin{pmatrix} \alpha_{g(i)} \\ \beta_{g(i)} \end{pmatrix}, \begin{pmatrix} \tau^2_a & \omega_{ab} \\ \omega_{ab} & \tau^2_b \end{pmatrix} \right)
\]

The covariance is the so-called G-matrix.

- \( G \) describes the population variance of the lines, i.e. the inter-individual variation.

PROC MIXED: random regression

PROC MIXED DATA=calcium;
CLASS grp girl; <--------------- Leave out visit
MODEL bmd=grp visit grp*visit / SOLUTION DDFM=SATTERTHWAITA;
RANDOM intercept visit / TYPE=UN SUBJECT=girl(grp) G VCORR;
RUN;

Individual intercepts and slopes are so-called random effects. They must be specified in a random-statement.

- Note that type=un refers to a unstructured specification of the G-matrix. If it is omitted, we may experience convergence problems and sometimes totally incomprehensible results.
- Option g asks that the estimated G-matrix be printed.
- ...
Output from random regression

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN(1,1)</td>
<td>girl(grp)</td>
<td>0.004105</td>
</tr>
<tr>
<td>UN(2,1)</td>
<td>girl(grp)</td>
<td>3.733E-6</td>
</tr>
<tr>
<td>UN(2,2)</td>
<td>girl(grp)</td>
<td>0.000048</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>0.000125</td>
</tr>
</tbody>
</table>

Fit Statistics

-2 Res Log Likelihood: -2341.6

Estimated G Matrix

<table>
<thead>
<tr>
<th>Row Effect</th>
<th>grp</th>
<th>girl</th>
<th>Col1</th>
<th>Col2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intercept</td>
<td>C</td>
<td>101</td>
<td>0.004105</td>
</tr>
<tr>
<td>2</td>
<td>time</td>
<td>C</td>
<td>101</td>
<td>3.733E-6</td>
</tr>
</tbody>
</table>

Note: The intercept refers to visit=0.

Implied covariance

The random regression model implies a particular covariance-structure:

\[
\text{Cov}(Y_{ij}, Y_{ik}) = \text{Cov}(A_i + B_i t_j + \varepsilon_{ij}, A_i + B_i t_k + \varepsilon_{ik}) \\
= \text{Var}(A_i) + (t_j + t_k)\text{Cov}(B_i, A_i) + t_j t_k \text{Var}(B_i) \\
= \tau_a^2 + (t_j + t_k)\omega_{ab} + t_j t_k \tau_b^2
\]

Does this fit the data well?

Random regression vs UN

\[
-2 \log Q = 2346.3 - 2341.6 \\
= 4.7 \sim \chi^2(15 - 3) = \chi^2(12) \Rightarrow P = 0.97
\]

We find an extra increase in BMD of 0.0045 (0.0016) g/cm\(^3\) per half year, when giving calcium supplement.

Implied covariance

▶ Option vcorr asks that the implied covariance- and correlation-matrices be printed.

Estimated V Matrix for girl(grp) 101 C

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
<th>Col5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.004285</td>
<td>0.004211</td>
<td>0.004263</td>
<td>0.004314</td>
<td>0.004366</td>
</tr>
<tr>
<td>2</td>
<td>0.004211</td>
<td>0.004435</td>
<td>0.004410</td>
<td>0.004509</td>
<td>0.004608</td>
</tr>
<tr>
<td>3</td>
<td>0.004263</td>
<td>0.004410</td>
<td>0.004681</td>
<td>0.004703</td>
<td>0.004850</td>
</tr>
<tr>
<td>4</td>
<td>0.004314</td>
<td>0.004509</td>
<td>0.004703</td>
<td>0.005022</td>
<td>0.005092</td>
</tr>
<tr>
<td>5</td>
<td>0.004366</td>
<td>0.004608</td>
<td>0.004850</td>
<td>0.005092</td>
<td>0.005459</td>
</tr>
</tbody>
</table>

Estimated V Correlation Matrix for girl(grp) 101 C

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
<th>Col5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>0.9660</td>
<td>0.9518</td>
<td>0.9300</td>
<td>0.9027</td>
</tr>
<tr>
<td>2</td>
<td>0.9660</td>
<td>1.0000</td>
<td>0.9677</td>
<td>0.9553</td>
<td>0.9364</td>
</tr>
<tr>
<td>3</td>
<td>0.9518</td>
<td>0.9677</td>
<td>1.0000</td>
<td>0.9700</td>
<td>0.9594</td>
</tr>
<tr>
<td>4</td>
<td>0.9300</td>
<td>0.9553</td>
<td>0.9700</td>
<td>1.0000</td>
<td>0.9725</td>
</tr>
<tr>
<td>5</td>
<td>0.9027</td>
<td>0.9364</td>
<td>0.9594</td>
<td>0.9725</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Implied covariance

We find an extra increase in BMD of 0.0045 (0.0016) g/cm\(^3\) per half year, when giving calcium supplement.
Non-equidistant time points

As it always happens:

- The girls are only seen approximately twice a year.
- Perhaps we get better estimates of the slopes when replacing visit with the actual age of the girl.
- But then an unstructured covariance cannot be estimated from the data ...

Some other covariance structures will still be possible, e.g. The non-equidistant analogue to the autoregressive structure ... and the random regression model.

Non-equidistant observations

In case subjects are measured at individual or otherwise non-equally spaced time points only a limited number of stationary covariance pattern models are available:

- The variance is constant over time.
- The correlation depend only on the time-distance between the observations.

Some other covariance structures will still be possible, e.g.

The non-equidistant analogue to the autoregressive structure ...

<table>
<thead>
<tr>
<th>proc mixed</th>
<th>Cov(Y_{ij}, Y_{ik})</th>
<th>no. param</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>σ^2[I{j = k} + ρ \cdot I{j ≠ k}]</td>
<td>2</td>
</tr>
<tr>
<td>SP(POW)(ctime)</td>
<td>σ^2ρ</td>
<td>t_{ij}−t_{ik}</td>
</tr>
<tr>
<td>SP(GAU)(ctime)</td>
<td>σ^2e^{−</td>
<td>t_{ij}−t_{ik}</td>
</tr>
<tr>
<td>SP(LIN)(ctime)</td>
<td>σ^2(1 −ρ</td>
<td>t_{ik}−t_{ij}</td>
</tr>
</tbody>
</table>

The ctime-variable must be a numerical variable in SAS.

Random regression, using actual age

Random regression: covariate age

- move intercept to 11 years (∼ age at first visit)
- \( y_{ij} = A_i + B_i(\text{age}_{ij} − 11) + ε_{ij} \)

In this model, we quantify the effect of a calcium supplement to 0.0089 (0.0031) g/cm\(^3\) per year.
### Results from random regression

<table>
<thead>
<tr>
<th>Group</th>
<th>Level at age 11</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>0.8667 (0.0087)</td>
<td>0.0453 (0.0022)</td>
</tr>
<tr>
<td>C</td>
<td>0.8778 (0.0088)</td>
<td>0.0542 (0.0022)</td>
</tr>
</tbody>
</table>

Difference: 0.0111 (0.0124) 0.0089 (0.0031)

P 0.37 0.0048

I.e. **steeper slopes** than when visit was used as the time-variable.

- Due to **quantification** (per year vs per 1/2 year)
- Due to **bias reduction** (visit is a proxy for age, and measurement error in the independent variable causes bias towards the null)

### Tests of treatment effect

**Comparison of estimates for different covariance structures:**

<table>
<thead>
<tr>
<th>Covariance structure</th>
<th>−2 log L</th>
<th>Cov.par.</th>
<th>Difference in slopes</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independence</td>
<td>-1245.0</td>
<td>1</td>
<td>0.0094 (0.0086)</td>
<td>0.27</td>
</tr>
<tr>
<td>Power (Autoregressive)</td>
<td>-2372.0</td>
<td>2</td>
<td>0.0094 (0.0032)</td>
<td>0.0038</td>
</tr>
<tr>
<td>Random Regression</td>
<td>-2350.1</td>
<td>4</td>
<td>0.0089 (0.0031)</td>
<td>0.0048</td>
</tr>
</tbody>
</table>

- **Confidence intervals and tests depend on the covariance!**

### Outline

- **Longitudinal analysis**
- **Models for the mean**
- **Covariance pattern models**
- **Random regression**
- **Baseline adjustments**

### Baseline measurements

In randomized clinical trials, the first measurement is often a **baseline measurement**.

- The group means **must** be equal at baseline!

When comparing the groups at baseline we can expect:

- A **non-significant differences** with 95% probability.
- A **type I error** occurring with 5% probability.

How is the overall test for treatment effect affected when mean at baseline is allowed to depend on the groups?
Hypothetical comparison of two treatment groups

What happens if we ignore the baseline problem?

The non-existing difference at baseline makes the overall treatment effect appear smaller. Thus, the power of the test is reduced.

So should we leave out the baseline measurement?

We lose information about change over time and again the power of the test of treatment effect is reduced.

Classical approaches for handling baseline


Three possibilities: 1. End point, 2. Change, 3. ANCOVA
   1. Discard baseline, works well when correlation is small
   2. Subtract baseline, works well when correlation is large
   3. Condition on baseline, using it as covariate, always works.

Conclusion: ANCOVA is most efficient.

Does this mean baseline should always be adjusted for?

Why ANCOVA is superior

For simplicity assume a two-group comparison, treatment vs placebo, with only one follow-up measurement \((t_1 = 0, t_2 = 1)\),

\[
Y_{ij} = \beta_1 + \beta_2 \cdot t_j + \beta_3 \cdot I\{\text{treat}\} \cdot t_j + \varepsilon_{ij}
\]

\[
\begin{pmatrix}
\varepsilon_{i1} \\
\varepsilon_{i2}
\end{pmatrix}
\sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2 \rho \\ \sigma^2 \rho & \sigma^2 \end{pmatrix}\right)
\]

- with variance that do not change with time

Implied residual variances for the three models.

1. \(\text{Var}(Y_2) = \sigma^2\)
2. \(\text{Var}(Y_2 - Y_1) = 2\sigma^2(1 - \rho^2)\)
3. \(\text{Var}(Y_2 | Y_1) = \sigma^2(1 - \rho^2)\)
ANCOVA with multiple times of follow-up

Different effects of baseline at different time points due to stronger correlation between baseline and early follow up.

- The model should include a baseline*time interaction!

```plaintext
proc sort data=calcium; by girl visit; run;
data base; set calcium; if visit EQ 1; bmd0=bmd; keep girl bmd0; run;
data fup; set calcium; if visit GE 2; run;
data calcium0; merge fup base; by girl; run;
proc mixed data=calcium0;
class grp girl visit;
model bmd=visit bmd0*visit grp*visit / ddfm=satterth s;
repeated visit / type=un subject=girl(grp) r rcorr;
run;
```

Calcium: ANCOVA

The residual covariance changes since baseline explains a good deal of variance in the data.

Estimated R Matrix for girl(grp) 101 C

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000284</td>
<td>0.000292</td>
<td>0.000312</td>
<td>0.000332</td>
</tr>
<tr>
<td>2</td>
<td>0.000292</td>
<td>0.000571</td>
<td>0.000573</td>
<td>0.000565</td>
</tr>
<tr>
<td>3</td>
<td>0.000312</td>
<td>0.000573</td>
<td>0.000778</td>
<td>0.000746</td>
</tr>
<tr>
<td>4</td>
<td>0.000332</td>
<td>0.000565</td>
<td>0.000746</td>
<td>0.000953</td>
</tr>
</tbody>
</table>

Estimated R Correlation Matrix for girl(grp) 101 C

<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>0.7262</td>
<td>0.6631</td>
<td>0.6388</td>
</tr>
<tr>
<td>2</td>
<td>0.7262</td>
<td>1.0000</td>
<td>0.8591</td>
<td>0.7649</td>
</tr>
<tr>
<td>3</td>
<td>0.6631</td>
<td>0.8591</td>
<td>1.0000</td>
<td>0.8655</td>
</tr>
<tr>
<td>4</td>
<td>0.6388</td>
<td>0.7649</td>
<td>0.8655</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Solution for Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>grp visit</th>
<th>Estimate</th>
<th>Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>0.08618</td>
<td>0.04472</td>
<td>99.6</td>
<td>1.93</td>
<td>0.0568</td>
<td></td>
</tr>
<tr>
<td>visit 2</td>
<td></td>
<td>-0.1197</td>
<td>0.03562</td>
<td>93</td>
<td>-3.37</td>
<td>0.0011</td>
<td></td>
</tr>
<tr>
<td>visit 3</td>
<td></td>
<td>-0.09062</td>
<td>0.02997</td>
<td>90.3</td>
<td>-3.02</td>
<td>0.0032</td>
<td></td>
</tr>
<tr>
<td>visit 4</td>
<td></td>
<td>-0.08082</td>
<td>0.02344</td>
<td>88.4</td>
<td>-3.45</td>
<td>0.0009</td>
<td></td>
</tr>
<tr>
<td>visit 5</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>bmd0*visit 2</td>
<td></td>
<td>1.0616</td>
<td>0.02684</td>
<td>102</td>
<td>39.55</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>bmd0*visit 3</td>
<td></td>
<td>1.0559</td>
<td>0.03862</td>
<td>100</td>
<td>27.34</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>bmd0*visit 4</td>
<td></td>
<td>1.0749</td>
<td>0.04580</td>
<td>101</td>
<td>23.47</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>bmd0*visit 5</td>
<td></td>
<td>1.0010</td>
<td>0.05116</td>
<td>99.6</td>
<td>19.67</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>grp*visit C2</td>
<td></td>
<td>0.006329</td>
<td>0.003293</td>
<td>102</td>
<td>1.92</td>
<td>0.0574</td>
<td></td>
</tr>
<tr>
<td>grp*visit C3</td>
<td></td>
<td>0.01172</td>
<td>0.004746</td>
<td>101</td>
<td>2.47</td>
<td>0.0152</td>
<td></td>
</tr>
<tr>
<td>grp*visit C4</td>
<td></td>
<td>0.01172</td>
<td>0.005592</td>
<td>99.9</td>
<td>2.10</td>
<td>0.0387</td>
<td></td>
</tr>
<tr>
<td>grp*visit C5</td>
<td></td>
<td>0.01896</td>
<td>0.006246</td>
<td>98.6</td>
<td>3.04</td>
<td>0.0031</td>
<td></td>
</tr>
</tbody>
</table>

Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>visit 3</td>
<td>93.2</td>
<td>6.64</td>
<td>0.0013</td>
<td></td>
</tr>
<tr>
<td>bmd0*visit 4</td>
<td>96.9</td>
<td>402.43</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>grp*visit 4</td>
<td>96.1</td>
<td>2.85</td>
<td>0.0280</td>
<td></td>
</tr>
</tbody>
</table>

Calcium: ANCOVA

Solution for Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>visit 2</th>
<th>visit 3</th>
<th>visit 4</th>
<th>visit 5</th>
<th>bmd0*visit 2</th>
<th>bmd0*visit 3</th>
<th>bmd0*visit 4</th>
<th>bmd0*visit 5</th>
<th>grp*visit C2</th>
<th>grp*visit C3</th>
<th>grp*visit C4</th>
<th>grp*visit C5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
<td>-0.1197</td>
<td>-0.09062</td>
<td>-0.08082</td>
<td>0</td>
<td>1.0616</td>
<td>1.0559</td>
<td>1.0749</td>
<td>1.0010</td>
<td>0.006329</td>
<td>0.01172</td>
<td>0.01172</td>
<td>0.01896</td>
</tr>
<tr>
<td>Error</td>
<td>0.03562</td>
<td>0.02997</td>
<td>0.02344</td>
<td>.</td>
<td>0.02684</td>
<td>0.03862</td>
<td>0.04580</td>
<td>0.05116</td>
<td>0.003293</td>
<td>0.004746</td>
<td>0.005592</td>
<td>0.006246</td>
</tr>
<tr>
<td>DF</td>
<td>93</td>
<td>90.3</td>
<td>88.4</td>
<td>.</td>
<td>102</td>
<td>100</td>
<td>101</td>
<td>99.6</td>
<td>102</td>
<td>101</td>
<td>99.9</td>
<td>100</td>
</tr>
<tr>
<td>t Value</td>
<td>-3.37</td>
<td>-3.02</td>
<td>-3.45</td>
<td>.</td>
<td>39.55</td>
<td>27.34</td>
<td>23.47</td>
<td>19.67</td>
<td>1.92</td>
<td>2.47</td>
<td>2.10</td>
<td>3.04</td>
</tr>
<tr>
<td>Pr &gt;</td>
<td>t</td>
<td></td>
<td>0.0011</td>
<td>0.0032</td>
<td>0.0009</td>
<td>.</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>0.0574</td>
<td>0.0152</td>
<td>0.0387</td>
<td>0.0031</td>
</tr>
</tbody>
</table>

The fourth option

Constrained linear mixed model (cLMM):

- Analysis of response profiles.
- include baseline as a response, but ...
- assume identical group means at baseline.


**Conclusion:** cLMM is the better option.

- about 1% higher power than ANCOVA with no missing data.
- no bias and higher power with missing data (MAR).

But: ANCOVA has a computational advantage, so ...
Power

Simulated 2 groups each of size \( n = 30 \)

\[ Y_{it} = \beta_1 + \beta_2 t + \beta_3 X_{it} + \varepsilon_{it} \]

\( i = 1, \ldots, n, \ t = 0, 1 \)

\( \varepsilon_i \)'s bivariate normal with variance \( \sigma^2 \) and correlation \( \rho \)

\[ \text{MIREDIF} = \beta_3 / \sigma. \]

For \( \rho = 0.6 \)

cLMM parametrisation

<table>
<thead>
<tr>
<th>( \text{Treatment} = P )</th>
<th>( \text{Treatment} = C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v=1 )</td>
<td>( \beta_1 )</td>
</tr>
<tr>
<td>( v=2 )</td>
<td>( \beta_1 + \beta_2 )</td>
</tr>
<tr>
<td>( v=3 )</td>
<td>( \beta_1 + \beta_3 )</td>
</tr>
<tr>
<td>( v=4 )</td>
<td>( \beta_1 + \beta_4 )</td>
</tr>
<tr>
<td>( v=5 )</td>
<td>( \beta_1 + \beta_5 )</td>
</tr>
</tbody>
</table>

- Intercept.
- Time effect in placebo-group
- Difference between groups at baseline = 0!
- Interactions (difference in time-effects)

Example: Calcium

In SAS we write:

data calciumx;
set calcium;
grpvisit = 'P1-S1';
if grp EQ 'C' and visit EQ 2 then grpvisit='C2';
if grp EQ 'C' and visit EQ 3 then grpvisit='C3';
if grp EQ 'C' and visit EQ 4 then grpvisit='C4';
if grp EQ 'C' and visit EQ 5 then grpvisit='C5';
run;

proc mixed data=calciumx;
class grp girl visit grpvisit;
model bmd = visit grpvisit / ddfm=satterth solution outpm=fit_clmm;
repeated visit / type=un subject=girl(grp) r rcorr;
run;

Calcium: cLMM output

<p>| Estimated R Matrix for girl(grp) 101 C |</p>
<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
<th>Col5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.003943</td>
<td>0.004186</td>
<td>0.004163</td>
<td>0.004238</td>
<td>0.003947</td>
</tr>
<tr>
<td>2</td>
<td>0.004186</td>
<td>0.004725</td>
<td>0.004709</td>
<td>0.004808</td>
<td>0.004519</td>
</tr>
<tr>
<td>3</td>
<td>0.004163</td>
<td>0.004709</td>
<td>0.004961</td>
<td>0.005042</td>
<td>0.004726</td>
</tr>
<tr>
<td>4</td>
<td>0.004238</td>
<td>0.004808</td>
<td>0.005042</td>
<td>0.005326</td>
<td>0.004980</td>
</tr>
<tr>
<td>5</td>
<td>0.003947</td>
<td>0.004519</td>
<td>0.004726</td>
<td>0.004980</td>
<td>0.004894</td>
</tr>
</tbody>
</table>

<p>| Estimated R Correlation Matrix for girl(grp) 101 C |</p>
<table>
<thead>
<tr>
<th>Row</th>
<th>Col1</th>
<th>Col2</th>
<th>Col3</th>
<th>Col4</th>
<th>Col5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>0.9698</td>
<td>0.9413</td>
<td>0.9249</td>
<td>0.8985</td>
</tr>
<tr>
<td>2</td>
<td>0.9698</td>
<td>1.0000</td>
<td>0.9727</td>
<td>0.9585</td>
<td>0.9398</td>
</tr>
<tr>
<td>3</td>
<td>0.9413</td>
<td>0.9727</td>
<td>1.0000</td>
<td>0.9809</td>
<td>0.9591</td>
</tr>
<tr>
<td>4</td>
<td>0.9249</td>
<td>0.9585</td>
<td>0.9809</td>
<td>1.0000</td>
<td>0.9755</td>
</tr>
<tr>
<td>5</td>
<td>0.8985</td>
<td>0.9398</td>
<td>0.9591</td>
<td>0.9755</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
Calcium: cLMM output

Solution for Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>grpvisit</th>
<th>visit</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>t</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.9622</td>
<td>0.007352</td>
<td>137</td>
<td>130.88</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit 1</td>
<td>-0.08706</td>
<td>0.004333</td>
<td>98.5</td>
<td>-20.09</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit 2</td>
<td>-0.06669</td>
<td>0.003434</td>
<td>93.8</td>
<td>-19.42</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit 3</td>
<td>-0.04261</td>
<td>0.002887</td>
<td>93</td>
<td>-14.76</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit 4</td>
<td>-0.01614</td>
<td>0.002289</td>
<td>93.3</td>
<td>-7.05</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grpvisit C2</td>
<td>0.006329</td>
<td>0.003272</td>
<td>103</td>
<td>1.93</td>
<td>0.0558</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grpvisit C3</td>
<td>0.01172</td>
<td>0.004709</td>
<td>101</td>
<td>2.49</td>
<td>0.0144</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grpvisit C4</td>
<td>0.01172</td>
<td>0.005548</td>
<td>101</td>
<td>2.11</td>
<td>0.0372</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grpvisit C5</td>
<td>0.01896</td>
<td>0.006195</td>
<td>99.4</td>
<td>3.06</td>
<td>0.0028</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num</th>
<th>Den</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>visit 4</td>
<td>97.9</td>
<td>111.12</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>grpvisit 4</td>
<td>96.9</td>
<td>2.90</td>
<td>0.0059</td>
<td></td>
</tr>
</tbody>
</table>

Interpretations

**ANCOVA and cLMMs estimate the treatment effect with similar accuracy (assuming an analysis of response profiles model)**

- P-values and parameter estimates are similar.
- But the two approaches are **not equivalent!**

Model parameters have different interpretations.

- Estimates in cLMM are population parameters.
- Estimates in ANCOVA are **conditional** on subjects having the same baseline response.

Baseline in observational studies

**Compare the outcomes for individuals from different groups (e.g. gender or illness groups):**

- The groups are likely to differ in many respects . . . **including the baseline outcome value!**
- Differences in response profiles may be due to many factors, and quantifications will depend on which of these are factors included in the model.
- Adjust for the covariates that are sensible in the context.

Is the baseline measurement a **sensible** covariate?

Fitzmaurice et al. (2011)[Section 5.6]:

For example, in an observational study examining gender differences in weight gain of infants between 12 months (baseline) and 24 months (...) At baseline boys are on average 1 1/2 pounds heavier than girls, but there is no evidence of a gender effect on the 12 month change in body weight, with boys and girls both gaining approximately 5 1/4 pound. In contrast the analysis of covariance of the same data reveals a discernible gender effect with boys showing more weight gain than girls.

(...) the analysis of covariance is directed at the conditional question of whether boys are expected to gain more weight than girls given that they have the same initial weight at 12 months. (...) The reasoning is that if a boy and girl have the same initial weight at 12 months, then there are two possibilities: (1) the girl is initially overweight and is expected to gain less weight or (2) the boy is initially underweight and is expected to gain more weight over the 12 months. We advise readers to employ the analysis of covariance approach in longitudinal settings only if the approach and its implications are fully understood.