Correlated measurements, terminology

Cluster design (previous lecture):
Same outcome (response) measured on all individuals in a number of families/villages/school classes

Repeated measurements (this + previous lecture):
Same outcome (response) measured in different situations (or at different spots) on the same individual.

Longitudinal measurements (this lecture):
Same outcome (response) measured consecutively over time for each individual.

Multivariate outcome (not treated):
Several outcomes (responses) for each individual, e.g. a number of hormone measurements that we want to study simultaneously.

Plan for today

Longitudinal designs
▶ Presenting data
▶ Advantages and drawbacks
▶ Common mistakes

Comparison of models
▶ Analysis of summary statistics
▶ Correlation structures
▶ Random regression
▶ Baseline effects

Analysis with SAS proc mixed.

Outline

The longitudinal design

Covariance structures

Random regression

Baseline effects

Summary of mixed models
Typical longitudinal design

Two or more groups of subjects.
- Randomized to different treatments
- Or healthy and ill receiving same treatment

Repeated measurements of the same quantity for each subject.

On a proper time scale:
- time since baseline / duration of treatment
- age
- cumulative dose of drug

Purpose of the investigation

Description of time course
- Do we see a change over time?
- Linear or curved?
- Same pattern for all groups?

Finding differences between the groups/treatments
- Same difference for all time points?
- Difference in level, or trend?

Traditional presentation of data

Aspirin absorption for healthy and ill subjects.

Problems with the traditional presentation of data

Comparison of groups for each time point separately:
- Is inefficient.
- Has a high risk of leading to chance significance.
- Interpretation may be difficult.

Average curves:
- Need not be representative.
- May hide important structures.
- Give no indication of the variation in the time profiles.
- Cannot be used for evaluation of changes over time because they do not show the pairing.
Individual time profiles

Always make a picture of individual time profiles!

Do not average over individual profiles, unless their shapes are identical, i.e. only shifts in level are seen between individuals.

Potential bias in mean value structure

Hypothetical example: Decline in 'health':

Individual time courses

Average curve

Individual characteristics

Examples of summary measures:

▶ Response for a selected time point, e.g. endpoint
▶ Average
▶ Slope, perhaps for a specific period
▶ Peak value
▶ Time to peak
▶ Area under the curve (AUC).
▶ A measure of cyclic behaviour.

Perform a traditional analysis with the summary measures as outcome (i.e. an ecological analysis).

Time to peak and peak values for aspirin absorption

Conclusion: P=0.02 for identity of peak values.
Repeated measurement designs

**Advantages:**
- Much more powerful in detecting time changes (the paired design situation)
- Informative about subject specific time courses. (Peaks and trends can be identified; In cross-sectional designs they cannot!)

**Drawbacks:**
- The traditional independence assumption is violated since repeated observations on the same individual are correlated.
- Statistical analyses are more complicated and less robust.

Traditional approaches for analysing time course

**Note:** Most of them requires a balanced design.

- Two-way anova or regression in subject and time may be used for describing time course in each group separately.
- Two-way anova or regression in group and time is wrong because it disregards the correlation within subjects.
- Three-way anova or regression in group, subject and time is impossible, since subjects are nested in groups.

Traditional approaches for analysing group differences

**Note:** Most of them requires a balanced design.

- Two-way anova or regression in group and time is wrong because it disregards the correlation within subjects.
- Three-way anova or regression in group, subject and time is impossible, since subjects are nested in groups.
- Comparisons of each specific time point may be ok, but cannot properly detect or quantify group differences in the overall time pattern.
- Comparison of time averages (or other summary measures) is often reasonable.

Outline

- The longitudinal design
- Covariance structures
- Random regression
- Baseline effects
- Summary of mixed models
Example: Calcium supplement to adolescent girls

**Study:** 112 11-year old girls randomised to either calcium supplement or placebo.

**Outcome:** BMD = bone mineral density (g/cm²) measured 5 times over 2 years at 6 month intervals.

**Scientific question:** Does calcium improve the rate of bone gain?

Two-level model

### Level 1

Variation

- within girls ($\sigma^2_W$)
- between girls ($\omega^2_B$)

**Covariates:** visit, grp*visit

$$y_{git} = \mu + \beta_g + \gamma_t + \delta_{gt} + \alpha_{gi} + \varepsilon_{git}$$

Random girl-levels, $a_{gi}$, correspond to having subject-specific intercepts, i.e. a so-called random intercept.

Analysis in SAS

```
proc mixed data=calcium;
class grp girl visit;
model bmd=grp visit grp*visit / ddfm=satterth s;
random girl(grp);
run;
```

Covariance Parameter Estimates (REML)

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIRL(GRP)</td>
<td>0.00443925</td>
</tr>
<tr>
<td>Residual</td>
<td>0.00025471</td>
</tr>
</tbody>
</table>

Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>NDF</th>
<th>DDF</th>
<th>Type III F</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRP</td>
<td>1</td>
<td>110</td>
<td>2.63</td>
<td>0.1078</td>
</tr>
<tr>
<td>VISIT</td>
<td>4</td>
<td>382</td>
<td>619.42</td>
<td>0.0001</td>
</tr>
<tr>
<td>GRP*VISIT</td>
<td>4</td>
<td>382</td>
<td>5.30</td>
<td>0.0004</td>
</tr>
</tbody>
</table>

Note: The specification `random girl(grp);` can be written in two other ways: random intercept / subject=girl(grp); or repeated visit / type=CS subject=girl(grp);

Estimates from the two-level model

<table>
<thead>
<tr>
<th>Effect</th>
<th>grp</th>
<th>visit</th>
<th>Estimate</th>
<th>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>C</td>
<td></td>
<td>0.9576</td>
<td>0.009131</td>
<td>122</td>
<td>104.87</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp</td>
<td>C</td>
<td></td>
<td>0.02951</td>
<td>0.01304</td>
<td>122</td>
<td>2.26</td>
<td>0.0254</td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp</td>
<td>P</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit</td>
<td>1</td>
<td>0</td>
<td>-0.08750</td>
<td>0.003101</td>
<td>382</td>
<td>-28.22</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit</td>
<td>2</td>
<td>-0.06748</td>
<td>0.003103</td>
<td>381</td>
<td>-21.75</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit</td>
<td>3</td>
<td>-0.04342</td>
<td>0.003117</td>
<td>381</td>
<td>-13.93</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit</td>
<td>4</td>
<td>-0.01619</td>
<td>0.003148</td>
<td>381</td>
<td>-6.14</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>visit</td>
<td>5</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>C</td>
<td>1</td>
<td>-0.01912</td>
<td>0.004445</td>
<td>382</td>
<td>-4.30</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>C</td>
<td>2</td>
<td>-0.01255</td>
<td>0.004448</td>
<td>381</td>
<td>-2.82</td>
<td>0.0050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>C</td>
<td>3</td>
<td>-0.00622</td>
<td>0.004480</td>
<td>381</td>
<td>-1.39</td>
<td>0.1661</td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>C</td>
<td>4</td>
<td>-0.00679</td>
<td>0.004517</td>
<td>381</td>
<td>-1.50</td>
<td>0.1337</td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>C</td>
<td>5</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>P</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>P</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>P</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>P</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grp*visit</td>
<td>P</td>
<td>5</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Visit no. 5 is the reference point.

The difference between groups tend to magnify with time.
Compound symmetry (CS)

The two-level model assumes that all measurements on the same individual are equally correlated:

\[
\text{Corr}(Y_{git_1}, Y_{git_2}) = \rho = \frac{\omega_B^2}{\omega_B^2 + \sigma^2_W}
\]

This is often inadequate as observations close to each other in time tend to be more correlated than observations further apart.

Note: Same model, many different names. Two-level model aka model with random subject levels aka model with random intercept aka model with compound symmetry correlation structure aka model with exchangeability correlation structure.

Output correlation matrix

Estimated R Correlation Matrix for girl(grp) 101 C

\[
\begin{pmatrix}
1 & \rho & \rho & \rho & \rho \\
\rho & 1 & \rho & \rho & \rho \\
\rho & \rho & 1 & \rho & \rho \\
\rho & \rho & \rho & 1 & \rho \\
\rho & \rho & \rho & \rho & 1
\end{pmatrix}
\]

Covariance matrix = total variation \times correlation matrix

Note: Distance in time is not taken into account; Observations are exchangeable.

Just checking: \( \hat{\rho} = \hat{\omega}_B^2 \hat{\omega}_B^2 + \hat{\sigma}_W^2 = 0.00443925 + 0.00023471 = 0.9498. \)

Compound symmetry analysis in SAS

proc mixed data=calcium;
  class grp girl visit;
  model bmd=grp visit grp*visit / ddfm=satterth outpredm=fit_cs;
  repeated visit / type=CS subject=girl(grp) rcorr;
run;

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>girl(grp)</td>
<td>0.004439</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>0.000235</td>
</tr>
</tbody>
</table>

Fit Statistics

-2 Res Log Likelihood = -2188.8

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>110</td>
<td>2.63</td>
<td>0.1078</td>
</tr>
<tr>
<td>visit</td>
<td>4</td>
<td>362</td>
<td>619.42</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>grp*visit</td>
<td>4</td>
<td>362</td>
<td>5.30</td>
<td>0.0004</td>
</tr>
</tbody>
</table>
Empirical correlation structure

Is compound symmetry reasonable?

<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>1.00000000</td>
<td>0.96987049</td>
<td>0.94138162</td>
<td>0.92499715</td>
<td>0.89865454</td>
</tr>
<tr>
<td>2</td>
<td>0.96987049</td>
<td>1.00000000</td>
<td>0.97270895</td>
<td>0.95852788</td>
<td>0.93987185</td>
</tr>
<tr>
<td>3</td>
<td>0.94138162</td>
<td>0.97270895</td>
<td>1.00000000</td>
<td>0.98090996</td>
<td>0.95919348</td>
</tr>
<tr>
<td>4</td>
<td>0.92499715</td>
<td>0.95852788</td>
<td>0.98090996</td>
<td>1.00000000</td>
<td>0.97553849</td>
</tr>
<tr>
<td>5</td>
<td>0.89865454</td>
<td>0.93987185</td>
<td>0.95919348</td>
<td>0.97553849</td>
<td>1.00000000</td>
</tr>
</tbody>
</table>

Other possibilities:
- Unstructured covariance (UN), →
- Patterned, e.g. autoregressive covariance (AR) → →
- Random regression → → →

Unstructured covariance (UN)

If we do not assume any specific structure for the covariance, we let it be arbitrary, i.e. unstructured.

Advantages:
- Does not force a wrong covariance structure on the data.
- Estimates the actual structure of the covariance.

Drawbacks:
- Uses quite a lot of parameters to describe the covariance structure, \( \frac{T(T+1)}{2} = 15 \) parameters for \( T = 5 \) visits. The result may therefore be unstable, in particular for small data sets.
- Can only be used with balanced data (i.e. all subjects have to be measured at identical times)

Unstructured covariance (UN) in SAS

```sas
proc mixed data=calcium;
  class grp girl visit;
  model bmd=grp visit grp*visit / ddfm=satterth outpredm=fit_un;
  repeated visit / type=UN subject=girl(grp) rcorr;
run;
```

Output from TYPE= UN model

```plaintext
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>1.00000000</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>1.000000</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>1.000000</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>1.000000</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>1.000000</td>
</tr>
</tbody>
</table>

Note: Same as the empirical correlation matrix.
Autoregressive covariance structure, AR(1)

In case of **equidistant times**, this specifies the following covariance matrix:

$$\sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 \\ \rho & 1 & \rho & \rho^2 & \rho^3 \\ \rho^2 & \rho & 1 & \rho & \rho^2 \\ \rho^3 & \rho^2 & \rho & 1 & \rho \\ \rho^4 & \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

I.e. the correlation decreases (in powers) with the distance between observations.

The **non-equidistant analogue** is: $\text{Corr}(Y_{git_1}, Y_{git_2}) = \rho|t_1-t_2|$.

Exponential decay

Correlation as function of distance between measurements.

Curves corresponds to $\rho = 0.1, \ldots, 0.9$.

Output from TYPE=AR(1) model

```
proc mixed data=calcium;
  class grp girl visit;
  model bmd=grp visit grp*visit / ddfm=satterth outpredm=fit_ar1;
  repeated visit / type=AR(1) subject=girl(grp) rcorr;
run;
```

```
Fit Statistics
-2 Res Log Likelihood -2318.6 <-----------used later

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>0.1005</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>
```

```
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>
```

```
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>Col15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</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>
```
Testing covariance structures

Likelihood ratio test:

- Good models have large values of likelihood \((L)\) and small values of deviance \((-2 \log L)\) which is outputted from SAS as \(-2\) Res Log Likelihood.
- Compute change in deviance \((\Delta = -2 \log Q)\) and compare to the \(\chi^2\)-distribution with \(df = \text{change in no. parameters}\).

Comparison of CS to UN:

\[-2 \log Q = 2346.3 - 2188.8 = 157.5 \sim \chi^2(15 - 2).\]

\(P < 0.0001,\) thus compound symmetry is not suitable.

Comparison of AR(1) to UN:

\[-2 \log Q = 2346.3 - 2318.6 = 27.7 \sim \chi^2(15 - 2).\]

\(P = 0.010,\) better but still not good enough.

Combination of CS and AR(1)

Neither of CS and AR(1) fits the data well enough - perhaps a combination will do?

Combined covariance matrix in case of equidistant times:

\[
\begin{pmatrix}
\omega^2 + \sigma^2 & \omega^2 + \sigma^2 \rho & \omega^2 + \sigma^2 \rho^2 & \omega^2 + \sigma^2 \rho^3 & \omega^2 + \sigma^2 \rho^4 \\
\omega^2 + \sigma^2 \rho & \omega^2 + \sigma^2 & \omega^2 + \sigma^2 \rho & \omega^2 + \sigma^2 \rho^2 & \omega^2 + \sigma^2 \rho^3 \\
\omega^2 + \sigma^2 \rho^2 & \omega^2 + \sigma^2 \rho & \omega^2 + \sigma^2 & \omega^2 + \sigma^2 \rho & \omega^2 + \sigma^2 \rho^2 \\
\omega^2 + \sigma^2 \rho^3 & \omega^2 + \sigma^2 \rho^2 & \omega^2 + \sigma^2 \rho & \omega^2 + \sigma^2 & \omega^2 + \sigma^2 \\
\omega^2 + \sigma^2 \rho^4 & \omega^2 + \sigma^2 \rho^3 & \omega^2 + \sigma^2 \rho^2 & \omega^2 + \sigma^2 \rho & \omega^2 + \sigma^2 \\
\end{pmatrix}
\]

Likelihoods in proc mixed

There are two different forms of likelihood:

- Default output is the REML-likelihood, where the mean value structure has been 'eliminated'.
- The full likelihood may be obtained using the option proc mixed method=ML;

For comparison of covariance structures:

- Use either of the two likelihoods.

For comparison of mean value structures:

- Use only the full likelihood!

Combination of CS and AR(1) in SAS

```sas
proc mixed data=calcium;
class grp girl visit;
model bmd=grp visit grp*visit / ddfm=satterth outpredm=fit_ar1;
random intercept / subject=girl(grp) g;
repeated visit / type=AR(1) subject=girl(grp) rcorr;
run;
```

Fit Statistics

<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>0.1055</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>
```
Output from CS + AR(1) model

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>girl(grp)</td>
<td>0</td>
</tr>
<tr>
<td>AR(1)</td>
<td>girl(grp)</td>
<td>0.9708</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>0.004413</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>
<th>Col5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</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>

Comparison of CS + AR(1) to UN:

\[-2 \log Q = 2346.3 - 2318.6 = 27.7 \sim \chi^2_{12}, \quad P = 0.006.\]

Hence, no improvement on AR(1)!

Test of no interaction

Test of grp*visit for various choices of covariance structure:

<table>
<thead>
<tr>
<th>Covariance structure</th>
<th>Test statistic ~ distribution</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independence</td>
<td>0.35 ~ F(4,491)</td>
<td>0.84</td>
</tr>
<tr>
<td>CS</td>
<td>5.30 ~ F(4,382)</td>
<td>0.0004</td>
</tr>
<tr>
<td>AR(1)</td>
<td>2.86 ~ F(4,382)</td>
<td>0.023</td>
</tr>
<tr>
<td>AR(1)+CS</td>
<td>2.90 ~ F(4,205)</td>
<td>0.023</td>
</tr>
<tr>
<td>UN</td>
<td>2.72 ~ F(4,107)</td>
<td>0.034</td>
</tr>
</tbody>
</table>

Summary

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>AR(1) + CS</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>

The autoregressive structure is definitely better than CS, but not quite good enough...

Additional replicates

What, if we had had double or triple measurements at each visit?

Analyze averages:

- If we always have the same number of replicates, a correct and optimal approach is to average them.
- If the number of repetitions vary, analysis of averages may still be valid (depending on the reason for the unbalance), although no longer optimal.

Three-level model:

- Modify proc mixed to: random girl girl*visit;
Predicted mean time profiles

Profiles are almost identical for all choices of covariance structures (here for the unstructured covariance).

Test of linear time trend

Trick: Define time=visit, keep visit, and do not include time in the class-statement.

proc mixed data=calcium;
class grp girl visit;
model bmd=grp time grp*time visit grp*visit / ddfm=satterth;
repeated visit / type=UN subject=girl(grp);
run;

<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>110</td>
<td>0.36</td>
<td>0.5485</td>
</tr>
<tr>
<td>time</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>time*grp</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>visit</td>
<td>3</td>
<td>97.7</td>
<td>3.61</td>
<td>0.0160</td>
</tr>
<tr>
<td>grp*visit</td>
<td>3</td>
<td>97.7</td>
<td>1.03</td>
<td>0.3849</td>
</tr>
</tbody>
</table>

Omiting the insignificant interaction

proc mixed data=calcium;
class grp girl visit;
model bmd=grp time grp*time visit / s ddfm=satterth;
repeated visit / type=UN subject=girl(grp) r;
run;

<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>0.34</td>
<td>0.5629</td>
</tr>
<tr>
<td>time</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>time*grp</td>
<td>1</td>
<td>97.2</td>
<td>8.12</td>
<td>0.0054</td>
</tr>
<tr>
<td>visit</td>
<td>3</td>
<td>98.8</td>
<td>3.65</td>
<td>0.0151</td>
</tr>
</tbody>
</table>

There is some deviation from linearity (P=0.0151), which we ought to investigate further...
Individual growth rates

Maybe the girls have different growth rates?

**Model** for i’th girl in g’th group at visit no. t:

\[ y_{git} = a_{gi} + b_{gi} t + \varepsilon_{git}, \]

where \( \varepsilon_{git} \sim N(0, \sigma^2_W) \).

I.e. different intercepts and different slopes for each girl.

Summary measures from individual regressions

<table>
<thead>
<tr>
<th></th>
<th>Level at age 11</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>Dif</td>
<td>0.0118 (0.0123)</td>
<td>0.0039 (0.0019)</td>
</tr>
</tbody>
</table>

\[ P = 0.34 \quad P = 0.050 \]

Analysis of covariance

Comparing mean slope of individual regressions to traditional analysis of covariance (wrongfully ignoring correlation).

Naive ANCOVA: 0.0049(0.0042), \( P=0.25 \)

Individual slopes: 0.0039(0.0019), \( P=0.050 \)

Random regression

We let each individual (girl) have her own level \( a_{gi} \) and her own slope \( b_{gi} \) assuming that these individual 'parameters' follow a bivariate normal distribution:

\[
\begin{pmatrix}
  a_{gi} \\
  b_{gi}
\end{pmatrix}
\sim N_2
\left(
\begin{pmatrix}
  \alpha_g \\
  \beta_g
\end{pmatrix},
\begin{pmatrix}
  \tau^2_a & \omega \\
  \omega & \tau^2_b
\end{pmatrix}
\right)
\]

The population variation of the lines, (i.e. the inter-individual variation) is described by the covariance matrix:

\[
G =
\begin{pmatrix}
  \tau^2_a & \rho \tau_a \tau_b & \omega \\
  \rho \tau_a \tau_b & \tau^2_b & \omega \\
  \omega & \omega & \tau^2_b
\end{pmatrix}
\]
Random regression in SAS

```sas
proc mixed covtest data=calcium;
class grp girl;
model bmd=grp time time*grp / ddfm=satterth s;
random intercept time / type=un subject=girl(grp) g vcorr;
run;
```

Estimated G Matrix

<table>
<thead>
<tr>
<th>Row</th>
<th>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>
<td>3.733E-6</td>
</tr>
<tr>
<td>2</td>
<td>time</td>
<td>C</td>
<td>101</td>
<td>3.733E-6</td>
<td>0.000048</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>

Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Run</th>
<th>Den</th>
<th>DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>grp</td>
<td>1</td>
<td>110</td>
<td>0.33</td>
<td>0.5685</td>
<td></td>
</tr>
<tr>
<td>time</td>
<td>1</td>
<td>96.4</td>
<td>985.55</td>
<td>&lt;.0001</td>
<td></td>
</tr>
</tbody>
</table>

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

Note concerning proc mixed notation

It is necessary to use TYPE=UN in the random-statement to allow intercept and slope to be arbitrarily correlated.

▶ In this example, the correlation between intercept and slope is not that impressive $\hat{\rho} = 0.0084$ that is. (Note that the intercept is not completely out of range since it refers to visit=0).

The default option in the random-statement is TYPE=VC, which only specifies variance components with different variances.

If TYPE=UN is omitted, we may experience convergence problems and sometimes totally incomprehensible results.

Non-equidistant samples

Actually – as it always happens –

▶ The girls are only seen approximately twice a year.
▶ The actual dates are available as the variable `ctime`, in SAS (internal date representation denoting days since ....).

Furthermore,

▶ The girls were not precisely 11 years at the first visit.
▶ We ought to use age of the girls as covariate, but unfortunately these are not available.
▶ We approximate age=$(ctime-11475)/365.25+12$.

▶ Note, that this will mostly affect the intercept estimates!
Non-equidistant samples in SAS

We can no longer use the construction `type=UN`, but still the `random`-statement and `type=CS` in the `repeated`-statement.

A lot of other covariance structures will still be possible, e.g. the non-equidistant analogue to the autoregressive structure,

$$\text{Corr}(Y_{git_1}, Y_{git_2}) = \rho|t_1-t_2|$$

which is written as `TYPE=SP(POW)(ctime)` in the `repeated`-statement.

---

Random regression, using actual age

```sas
proc mixed covtest data=calcium;
class grp girl;
model bmd=grp age11 age11*grp / ddfm=satterth s outpm=predicted_mean;
random intercept age11 /
    type=un subject=girl(grp) g vcorr;
run;
```

**Solution for Fixed Effects**

| Effect    | grp | Estimate | Standard Error | DF | t Value | Pr > |t| |
|-----------|-----|----------|----------------|----|---------|------|---|
| Intercept |     | 0.8667   | 0.008688       | 110| 99.75   | <0.0001 |
| grp       | C   | 0.01113  | 0.01240        | 110| 0.90    | 0.3715  |
|            | P   | 0        |                |    |         |       |   |
| age11     | C   | 0.04529  | 0.002152       | 96 | 21.05   | <0.0001 |
|            | P   | 0        |                |    |         |       |   |
| age11*grp | C   | 0.008891 | 0.003076       | 96.6| 2.89    | 0.0048  |
|            | P   | 0        |                |    |         |       |   |

In this model, we quantify the effect of a calcium supplement to 0.0089 (0.0031) g per cm³ **per year**.

---

Comparison of slope estimates

...for different covariance structures:

<table>
<thead>
<tr>
<th>Covariance structure</th>
<th>Difference in slopes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independence</td>
<td>0.0094 (0.0086)</td>
<td>0.27</td>
</tr>
<tr>
<td>Compound symmetry</td>
<td>0.0089 (0.0020)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Exponential (Autoregressive)</td>
<td>0.0094 (0.0032)</td>
<td>0.0038</td>
</tr>
<tr>
<td>Random regression</td>
<td>0.0089 (0.0031)</td>
<td>0.0048</td>
</tr>
</tbody>
</table>

---

Random vs individual regression

We compare the results of the random regression to the analysis of summary measures: individual intercept and slope.

**Random regression:**

- **Level at age 11**
  - P: 0.8667 (0.0087)
  - C: 0.8778 (0.0088)
- **Slope**
  - P: 0.0453 (0.0022)
  - C: 0.0542 (0.0022)

**Dif**

- P: 0.0111 (0.0124)
- C: 0.0089 (0.0031)

**Individual regresions:**

- **Level at age 11**
  - P: 0.8697 (0.0086)
  - C: 0.8815 (0.0088)
- **Slope**
  - P: 0.0412 (0.0028)
  - C: 0.0488 (0.0028)

**Dif**

- P: 0.0118 (0.0123)
- C: 0.0076 (0.0038)

**P** = 0.37

**P** = 0.0048

**P** = 0.34

**P** = 0.050
Predicted values from random regression

It looks as if there is a difference right from the start (although we have previously seen this to be *insignificant*, $P=0.37$).

▶ **Do we need a baseline adjustment?**

---

Baseline effects

The first visit is a **baseline measurement** just after randomization was performed.

The two groups are **known to be equal at baseline**.

Including baseline measurements in the comparison of the groups

▶ May weaken the power for finding a possible difference between the treatments.
▶ May convert the treatment effect to an interaction →

---

Hypothetical comparison of two treatment groups

▶ **Truth:** Constant difference between the treatments.
▶ **Finding:** Interaction between time and treatment.
Systematic baseline differences

Even when randomization is properly conducted, dissimilarities may be present in small studies.

For ‘slowly varying’ outcomes (i.e. high degree of correlation), even a small difference may produce non-treatment related differences, i.e. bias.

Take care with interpretations; Any effect may be due to either group or baseline!

If randomization went wrong, baseline must be accounted for in the analysis.

Hypothetical comparison of two treatment groups

▶ Finding: Constant difference between treatments.

Approaches for handling individual baseline differences

Use follow-up data only (exclude baseline from analysis).
▶ most reasonable if correlation between repeated measurements is very low.

Subtract baseline from successive measurements.
▶ most reasonable if correlation between repeated measurements is very high.

Use baseline measurement as covariate.
▶ may be used for any degree of correlation.

Repeated measurements model accounts for differences between individuals and varying degrees of correlation!

Baseline correction in the calcium example

Including baseline as a (level 2) covariate will hardly change the difference between the slopes since they are within-individual quantities (coefficients for the age covariate).

A small change is expected because of the exclusion of visit 1 from the analysis, and because slope is correlated with level.

The difference between groups at fixed ages – e.g. endpoint age of 13 years – may be affected by the baseline correction.
Random regression

No baseline correction yet!

```
proc mixed covtest noclprint data=calcium;
class grp girl;
model bmd=grp age13 grp*age13 / ddfm=satterth s;
random intercept age13 / type=un subject=girl(grp) g;
run;
```

Solution for Fixed Effects

| Effect   | grp | Estimate | Error   | DF  | t Value | Pr > |t| |
|----------|-----|----------|---------|-----|---------|-------|
| Intercept|     | 0.9573   | 0.009819| 108 | 97.49   | <.0001|
| grp C    |     | 0.02891  | 0.01402 | 108 | 2.06    | 0.0416|
| age13    |     | 0.04529  | 0.002152| 96  | 21.05   | <.0001|
| age13*grp| C   | 0.008891 | 0.003076| 96.6| 2.89    | 0.0048|

Estimated gain at the age 13: 0.0289 (0.0140) g per cm$^3$

Including baseline as covariate

```
proc mixed covtest noclprint data=calcium; where visit>1;
class grp girl;
model bmd=baseline grp age13 grp*age13 / ddfm=satterth s;
random intercept age13 / type=un subject=girl(grp) g;
run;
```

Solution for Fixed Effects

| Effect   | grp | Estimate | Error   | DF  | t Value | Pr > |t| |
|----------|-----|----------|---------|-----|---------|-------|
| Intercept|     | 0.01825  | 0.02690 | 106 | 0.68    | 0.4989|
| baseline |     | 1.0797   | 0.03064 | 102 | 35.36   | <.0001|
| grp C    |     | 0.01728  | 0.006236| 101 | 2.77    | 0.0067|
| age13    |     | 0.04597  | 0.002287| 93.1| 20.11   | <.0001|
| age13*grp| C   | 0.007466 | 0.003277| 92.5| 2.28    | 0.0252|

Estimated gain at the age 13: 0.0173 (0.0062) g per cm$^3$

Follow-up visits only

```
proc mixed covtest noclprint data=calcium; where visit>1;
class grp girl;
model bmd=grp age13 grp*age13 / ddfm=satterth s;
random intercept age13 / type=un subject=girl(grp) g;
run;
```

Solution for Fixed Effects

| Effect   | grp | Estimate | Error   | DF  | t Value | Pr > |t| |
|----------|-----|----------|---------|-----|---------|-------|
| Intercept|     | 0.9574   | 0.009721| 102 | 98.49   | <.0001|
| grp C    |     | 0.02474  | 0.01383 | 102 | 1.79    | 0.0765|
| age13    |     | 0.04634  | 0.002288| 92.3| 20.25   | <.0001|
| age13*grp| C   | 0.007456 | 0.003277| 92.5| 2.28    | 0.0252|

Estimated gain at the age 13: 0.0247 (0.0138) g per cm$^3$

Summary

Estimated difference between groups at age 13:
- without baseline correction: 0.0289 (0.0140)
- based on follow up data only: 0.0247 (0.0138)
- with baseline as covariate: 0.0173 (0.0062)

Including baseline as a covariate explains some (but not all) of the difference between groups at age 13 and increases the precision of the estimated difference (standard error becomes smaller).

▶ The difference becomes more significant!
Only two time points: baseline and follow-up


52 patients with shoulder pain are randomized to either:
- Acupuncture (n=25)
- Placebo (n=27)

Pain is evaluated on a 100 point scale before and after treatment.
- High scores are good

Approaches for pain score analysis

The acupuncture group lies somewhat above placebo at baseline

Follow-up data only:
- We would expect the acupuncture group to be higher also after treatment. Therefore, a direct comparison of follow-up times is unreasonable (we see too big a difference).

Subtract baseline:
- Low baseline implies that the placebo group is expected to increase the most (regression to the mean). Therefore, a direct comparison of changes in pain score is unreasonable (we see too small a difference).

Use baseline as a covariate!

Development of pain, actual and hypothetical

Results on pain scores

<table>
<thead>
<tr>
<th>Type of analysis</th>
<th>Average pain score</th>
<th>Treatment effect difference (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>53.9 (14.0)</td>
<td>6.5</td>
<td>0.09</td>
</tr>
<tr>
<td>Type of analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>Average placebo (n=27) 62.3 (17.9)</td>
<td>17.3 (7.5; 27.1)</td>
<td>0.0008</td>
</tr>
<tr>
<td>Subtract*</td>
<td>Average acupuncture (n=25) 79.6 (17.1)</td>
<td>10.8 (2.3; 19.4)</td>
<td>0.014</td>
</tr>
<tr>
<td>Baseline as covariate</td>
<td></td>
<td>12.7 (4.1; 21.3)</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Outline

The longitudinal design

Covariance structures

Random regression

Baseline effects

Summary of mixed models

Specification of mixed models

Systematic variation:
- Between-individual E.g. treatment, age, baseline.
- Within-individual E.g. time, cumulative dose.

Random variation:
- Random effects, serial correlation, measurement error.
- Note: Interactions between systematic and random effects are always random

SAS, PROC MIXED

model
- describes the systematic part (fixed effects, mean value structure)

random
- describes the random effects

repeated
- describes the serial correlation

local
- adds an additional measurement error

The option ddfm=satterth

In balanced designs:
- The approximation has no effect.

In unbalanced designs (or balanced with missing observations):
- The approximation is necessary.

It may give rise to fractional degrees of freedom!

The computations may require a little more time, but in most cases this will not be noticeable.

When in doubt, use it!