

Repeated Measures Data

Data where subjects are measured repeatedly over time

- predetermined intervals (weekly)
- uncontrolled – variable intervals between measurements

Reasons for repeated measures

- ensure treatment is effective over a specified period
- monitor safety aspects of the treatment over a specified period
- measure length of time for treatment to become effective

Objectives of repeated measures analysis

- measure average treatment effect over time
- assess possible time*treatment interactions
- assess treatment response profile (e.g. Area Under Curve (AUC), tmax, cmax)
- identify possible covariance patterns in the repeated measurements
- determine an appropriate model to describe the time and measurement relationship.

Fixed effects approaches

Although it is possible to treat the patient/animal as a fixed effect in special situations, normally the patient/animal is treated as random. The book briefly covers a few of these situations in Chapter 6 and covers the advantages and disadvantages of these. Most of them require fixed time intervals and no missing data.

Mixed model approaches

Advantages of using mixed models with repeated measures data

- a single model can be used to estimate overall treatment effects and to estimate treatment effects at each time point.
- Missing data does not cause problems as long as they can be assumed missing at random.
- The covariance pattern of the repeated measurements can be determined and taken into account.

Covariance Pattern Models

Several ways to analyze repeated measures data using mixed models are available. The first one we will cover is the random effects model with animal effects treated as random. It allows for constant correlation between all observations on the same animal. However,

correlations between observations on the same animal are not always constant. Observations recorded on consecutive visits may be more highly correlated than observations measured on visits further apart. A **covariance pattern model** can be used to model this.

The basic structure of the covariance pattern model is

$$R = \begin{bmatrix} R_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & R_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & R_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & R_4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & R_5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & R_6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & R_7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_{10} \end{bmatrix}$$

Where the submatrix R_i covariance block correspond to the i th animal where the dimension of the submatrix correspond to the number of repeated measurements for animal i . The 0's are also block matrices that correspond between observations of different animals and assumes that the correlation between observations among animals are zero.

Covariance patterns

There are a large number of possible covariance patterns for each of the submatrices R_i available for mixed models. We will present several of the more commonly used patterns which are available in SAS.

Simple covariance patterns

For our example, we will include 3 animals, where Animal 1 had 3 visits, Animal 2 had 4 visits and Animal 3 had 2 visits, where the visits are at fixed times.

Independent covariance pattern

If the within-subjects correlation is zero, the covariance model is called independent and is the simplest model. The number of parameters that need to be estimated is equal to the number of visits and assumes that the visits are independent. This covariance structure is also referred to as “heterogeneous uncorrelated” because it allows for differing variances at each visit. For our example, the independent covariance pattern would be

$$R = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_1^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_2^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_3^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_4^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_1^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_2^2 \end{bmatrix}$$

General or unstructured covariance pattern

However, in most experiments there is more than likely some correlation present. If the correlation between each pair of visits (θ_{ij}) (within-subjects correlation) is allowed to differ and the variance for each time point (σ_i^2) is allowed to differ, then the covariance pattern is called general or unstructured. This is the most complex of the “simple correlation patterns”. The number of parameters that need to be estimated is $(N*(N+1))/2$. For example, if there are 4 visits ($N=4$), then the number of parameters would be $(4*5)/2=10$. The general covariance pattern for our example follows:

$$R = \begin{bmatrix} \sigma_1^2 & \theta_{12} & \theta_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta_{12} & \sigma_2^2 & \theta_{23} & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta_{13} & \theta_{23} & \sigma_3^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_1^2 & \theta_{12} & \theta_{13} & \theta_{14} & 0 & 0 \\ 0 & 0 & 0 & \theta_{12} & \sigma_2^2 & \theta_{23} & \theta_{24} & 0 & 0 \\ 0 & 0 & 0 & \theta_{13} & \theta_{23} & \sigma_3^2 & \theta_{34} & 0 & 0 \\ 0 & 0 & 0 & \theta_{14} & \theta_{24} & \theta_{34} & \sigma_4^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_1^2 & \theta_{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_{12} & \sigma_2^2 \end{bmatrix}$$

First-order autoregressive covariance pattern

When time periods are evenly spaced, the first-order autoregressive model is sometimes appropriate. This model is also referred to as AR(1). In this model, the variances are the same for all of the visits and the covariances decrease exponentially depending on their distance ($\theta_{ij} = \rho^{|i-j|}\sigma^2$). Note that the covariance is a function of the variance so the scalar variance can be pulled out of the matrix and that there are only two parameters that need to be estimated. For our example, the first-order autoregressive would be:

$$R = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho & 1 & \rho & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho^2 & \rho & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \rho & \rho^2 & \rho^3 & 0 & 0 \\ 0 & 0 & 0 & \rho & 1 & \rho & \rho^2 & 0 & 0 \\ 0 & 0 & 0 & \rho^2 & \rho & 1 & \rho & 0 & 0 \\ 0 & 0 & 0 & \rho^3 & \rho^2 & \rho & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \rho \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho & 1 \end{bmatrix}$$

Compound symmetry covariance pattern

If all of the covariances among the visits are thought to be the same, then the compound symmetry covariance pattern is used. This pattern does not make any assumption about the relationship between the variance and covariance and only has two parameters that need to be estimated. For our example, this would be:

$$R = \begin{bmatrix} \sigma^2 & \theta & \theta & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta & \sigma^2 & \theta & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta & \theta & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 & \theta & \theta & \theta & 0 & 0 \\ 0 & 0 & 0 & \theta & \sigma^2 & \theta & \theta & 0 & 0 \\ 0 & 0 & 0 & \theta & \theta & \sigma^2 & \theta & 0 & 0 \\ 0 & 0 & 0 & \theta & \theta & \theta & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & \theta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta & \sigma^2 \end{bmatrix}$$

Toeplitz covariance pattern

The Toeplitz covariance is similar to the first-order autoregressive covariance pattern in the way that the pairs separated by the same distance share the same correlations. But with the Toeplitz covariance pattern there is no known function relating the ρ values to the distance. This covariance pattern makes sense when visits are equally spaced. The number of parameters that need to be estimated is equal to the maximum number of visits, which is greater than the two parameters needed for AR(1). For our example, the Toeplitz covariance pattern would be

$$R = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho_1 & 1 & \rho_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho_2 & \rho_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \rho_1 & \rho_2 & \rho_3 & 0 & 0 \\ 0 & 0 & 0 & \rho_1 & 1 & \rho_1 & \rho_2 & 0 & 0 \\ 0 & 0 & 0 & \rho_2 & \rho_1 & 1 & \rho_1 & 0 & 0 \\ 0 & 0 & 0 & \rho_3 & \rho_2 & \rho_1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \rho_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_1 & 1 \end{bmatrix}$$

Heterogeneous variances among visits

The first-order autoregressive, compound symmetry and Toeplitz covariance patterns all assumed that the variances for all of the visits were the same. However, sometimes variability will differ between the visits for the trait being measured. In other words, the variances are heterogeneous. The independent and general covariance patterns allowed for heterogeneity among the variances. The AR(1), compound symmetry and Toeplitz covariance patterns can also be modified to allow for heterogeneous variances and are presented for our example.

Heterogeneous first-order autoregressive

The number of parameters needed for the heterogeneous AR(1) is the number of visits plus 1. In our example, with 4 visits, we have 5 parameters that need to be estimated.

$$R = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho^2\sigma_1\sigma_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 & \rho\sigma_2\sigma_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho^2\sigma_1\sigma_3 & \rho\sigma_2\sigma_3 & \sigma_3^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho^2\sigma_1\sigma_3 & \rho^3\sigma_1\sigma_4 & 0 & 0 \\ 0 & 0 & 0 & \rho\sigma_1\sigma_2 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho^2\sigma_2\sigma_4 & 0 & 0 \\ 0 & 0 & 0 & \rho^2\sigma_1\sigma_3 & \rho\sigma_2\sigma_3 & \sigma_3^2 & \rho\sigma_3\sigma_4 & 0 & 0 \\ 0 & 0 & 0 & \rho^3\sigma_1\sigma_4 & \rho^2\sigma_2\sigma_4 & \rho\sigma_3\sigma_4 & \sigma_4^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_1^2 & \rho\sigma_1\sigma_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Heterogeneous compound symmetry

The number of parameters needed for the heterogeneous compound symmetry is also the number of visits plus 1.

$$R = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 & \rho\sigma_2\sigma_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho\sigma_1\sigma_3 & \rho\sigma_2\sigma_3 & \sigma_3^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 & \rho\sigma_1\sigma_4 & 0 & 0 \\ 0 & 0 & 0 & \rho\sigma_1\sigma_2 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \rho\sigma_2\sigma_4 & 0 & 0 \\ 0 & 0 & 0 & \rho\sigma_1\sigma_3 & \rho\sigma_2\sigma_3 & \sigma_3^2 & \rho\sigma_3\sigma_4 & 0 & 0 \\ 0 & 0 & 0 & \rho\sigma_1\sigma_4 & \rho\sigma_2\sigma_4 & \rho\sigma_3\sigma_4 & \sigma_4^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_1^2 & \rho\sigma_1\sigma_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Heterogeneous Toeplitz

Just as with the homogeneous covariance patterns, the number of parameters that need to be estimated for the heterogeneous Toeplitz covariance pattern is greater than for heterogeneous AR(1) or compound symmetry. Again, this is because there is no known function relating the ρ values to the distance. For our example, the heterogeneous Toeplitz covariance pattern would be

$$R = \begin{bmatrix} \sigma_1^2 & \rho_1\sigma_1\sigma_2 & \rho_2\sigma_1\sigma_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho_1\sigma_1\sigma_2 & \sigma_2^2 & \rho_1\sigma_2\sigma_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho_2\sigma_1\sigma_3 & \rho_1\sigma_2\sigma_3 & \sigma_3^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_1^2 & \rho_1\sigma_1\sigma_2 & \rho_2\sigma_1\sigma_3 & \rho_3\sigma_1\sigma_4 & 0 & 0 \\ 0 & 0 & 0 & \rho_1\sigma_1\sigma_2 & \sigma_2^2 & \rho_1\sigma_2\sigma_3 & \rho_2\sigma_2\sigma_4 & 0 & 0 \\ 0 & 0 & 0 & \rho_2\sigma_1\sigma_3 & \rho_1\sigma_2\sigma_3 & \sigma_3^2 & \rho_1\sigma_3\sigma_4 & 0 & 0 \\ 0 & 0 & 0 & \rho_3\sigma_1\sigma_4 & \rho_2\sigma_2\sigma_4 & \rho_1\sigma_3\sigma_4 & \sigma_4^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_1^2 & \rho_1\sigma_1\sigma_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_1\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

First-order ante dependence covariance pattern

If visits are not equally spaced and the correlation structure changes over time, the AR(1) and Toeplitz covariance patterns no longer make the most sense. The first-order ante dependence covariance pattern ANTE(1) is a more general model that keeps the main features of the AR(1) and Toeplitz, but allows for unequal visit spacing and changes in the correlation structure over time.

$$R = \begin{bmatrix} \sigma_1^2 & \rho_1\sigma_1\sigma_2 & \rho_1\rho_2\sigma_1\sigma_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho_1\sigma_1\sigma_2 & \sigma_2^2 & \rho_2\sigma_2\sigma_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho_1\rho_2\sigma_1\sigma_3 & \rho_2\sigma_2\sigma_3 & \sigma_3^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_1^2 & \rho_1\sigma_1\sigma_2 & \rho_1\rho_2\sigma_1\sigma_3 & \rho_1\rho_2\rho_3\sigma_1\sigma_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_1\sigma_1\sigma_2 & \sigma_2^2 & \rho_2\sigma_2\sigma_3 & \rho_2\rho_3\sigma_2\sigma_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_1\rho_2\sigma_1\sigma_3 & \rho_2\sigma_2\sigma_3 & \sigma_3^2 & \rho_3\sigma_3\sigma_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_1\rho_2\rho_3\sigma_1\sigma_4 & \rho_2\rho_3\sigma_2\sigma_4 & \rho_3\sigma_3\sigma_4 & \sigma_4^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_1^2 & \rho_1\sigma_1\sigma_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_1\sigma_1\sigma_2 & \sigma_2^2 & 0 \end{bmatrix}$$

Note that the ANTE(1) covariance pattern

- permits the variance among observations to change over time
- permits the correlations to change over time (correlations between pairs of observations is the product of the correlations between adjacent times between observations)
- 2N-1 parameters estimated. For our example 2*4-1=7.

Separate covariance patterns for each treatment

Response to treatment may differ in amount of variability, so different sets of covariance parameters may be used for each treatment group. For our example, Animals 1 and 3 received Treatment A and Animal 2 received Treatment B. Previously, when we assumed homogeneous treatment variances, the number of parameters that needed to be estimated was two. Assuming heterogeneous treatment variances increases the number of parameters multiplicatively by the number of treatments. For our two treatment example, this would be 2*2=4 parameters. If the compound symmetry covariance structure is used, the structure we would now have would be

$$R = \begin{bmatrix} \sigma_A^2 & \theta_A & \theta_A & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta_A & \sigma_A^2 & \theta_A & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta_A & \theta_A & \sigma_A^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_B^2 & \theta_B & \theta_B & \theta_B & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_B & \sigma_B^2 & \theta_B & \theta_B & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_B & \theta_B & \sigma_B^2 & \theta_B & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_B & \theta_B & \theta_B & \sigma_B^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_A^2 & \theta_A & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_A & \sigma_A^2 & 0 \end{bmatrix}$$

In our example, we present only 3 animals. If we assume that there are additional animals in the study and that the maximum number of visits for the study is 4, for the

general or unstructured pattern, the number of parameters in our example would be $2*10=20$.

$$R = \begin{bmatrix} \sigma_{A,1}^2 & \theta_{A,12} & \theta_{A,13} & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta_{A,12} & \sigma_{A,2}^2 & \theta_{A,23} & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta_{A,13} & \theta_{A,23} & \sigma_{A,3}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{B,1}^2 & \theta_{B,12} & \theta_{B,13} & \theta_{B,14} & 0 & 0 \\ 0 & 0 & 0 & \theta_{B,12} & \sigma_{B,2}^2 & \theta_{B,23} & \theta_{B,24} & 0 & 0 \\ 0 & 0 & 0 & \theta_{B,13} & \theta_{B,23} & \sigma_{B,3}^2 & \theta_{B,34} & 0 & 0 \\ 0 & 0 & 0 & \theta_{B,14} & \theta_{B,24} & \theta_{B,34} & \sigma_{B,4}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{A,1}^2 & \theta_{A,12} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_{A,12} & \sigma_{A,2}^2 \end{bmatrix}$$

Banded covariances

With the exception of the independent structure, all of the above covariance pattern structures assumed that there was some correlation among visits. As the distance in visit number between two different visits increases, the correlation between the observations from the visits in most cases decreases. The correlations between widely separated observations is close to zero and banding the covariance matrices by setting correlations between these observations to zero may be appropriate. The advantage of using a banded covariance structure is that it reduces the number of parameters that need to be estimated. Banding is very useful in experiments where there are a large number of visits or measurement time points. For example, if in our example, we used a band size of 3 with a general covariance pattern, we would get the following, where the θ_{34} correlation parameter is assumed to be zero and is not estimated:

$$R = \begin{bmatrix} \sigma_1^2 & \theta_{12} & \theta_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta_{12} & \sigma_2^2 & \theta_{23} & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta_{13} & \theta_{23} & \sigma_3^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_1^2 & \theta_{12} & \theta_{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_{12} & \sigma_2^2 & \theta_{23} & \theta_{24} & 0 & 0 \\ 0 & 0 & 0 & \theta_{13} & \theta_{23} & \sigma_3^2 & \theta_{34} & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_{24} & \theta_{34} & \sigma_4^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_1^2 & \theta_{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_{12} & \sigma_2^2 \end{bmatrix}$$

Selecting an appropriate covariance pattern

There are many choices of covariance patterns for modeling repeated measures data. Choosing the most appropriate pattern is important in order to draw accurate conclusions.

If the covariance pattern chosen is too simple, then you risk increasing Type I error rate and underestimating standard errors. If the covariance pattern chosen is too complex, you sacrifice power and efficiency. As more covariance parameters are included, the chances of over fitting increase. There are tools available to help select a covariance pattern. First, there are graphical tools to help you visualize patterns of correlation between observations at different times. We will go through examples of graphical tools later. Secondly, there are measures of model fit, including **information criteria** and statistical comparison between models.

Information criteria

Information criteria measure the relative fit of competing covariance patterns. As more parameters are included in the model, the likelihood statistic is expected to become larger. There are two statistics based on the likelihood that make allowance for the number of covariance parameters fitted and can be used to compare two models which fit *the same fixed effects*. **Akaike's information criterion (AIC)**

$$\text{AIC} = \log(L) - q$$

Where q is the number of covariance parameters and $\log(L)$ is the log of the likelihood for the model. SAS calculates the AIC as $-2(\log(L)) + 2q$.

The second statistic is **Schwarz's information criterion (SIC)**

$$\text{SIC} = \log(L) - (q * \log(N - p)) / 2$$

Where p is the number of fixed effects, N is the number of observations and q is the number of covariance parameters. This is also referred to as the BIC (Bayesian information criterion) or SBC (Schwarz's Bayesian information criterion). SAS calculates the BIC as $-2(\log(L)) + (q * \log(N - p))$ where N is the number of subjects.

Models with larger values of AIC and SIC as presented in the text mean better fit. With SAS PROC MIXED, smaller means a better fit. Whether AIC or BIC is better is not clear.

Statistical comparison between models

The likelihood ratio test can also be used to compare models which fit the same fixed effects and whose covariance patterns are nested. Nesting is when the covariance pattern in the simpler model can be obtained by restricting some of the parameters in the more complex model. An example of this would be as follows:

$$R = \begin{bmatrix} \sigma^2 & \theta & \theta \\ \theta & \sigma^2 & \theta \\ \theta & \theta & \sigma^2 \end{bmatrix} \text{ compound symmetry pattern is nested within}$$

$$R = \begin{bmatrix} \sigma^2 & \theta_1 & \theta_2 \\ \theta_1 & \sigma^2 & \theta_1 \\ \theta_2 & \theta_1 & \sigma^2 \end{bmatrix} \text{ Toeplitz pattern, because a restriction on the Toeplitz pattern that}$$

$\theta_1 = \theta_2$ would be the same as the compound symmetry pattern.

However, the compound symmetry pattern is not nested within the AR(1)

$$R = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix} \text{ because the off diagonal elements in the AR(1) are a function of}$$

both the σ^2 and the ρ .

The likelihood ratio test statistic is

$2(\log(L_1) - \log(L_2)) \sim \chi_{df}^2$ where df = difference in number of covariance parameters fitted.

Which covariance patterns to consider?

A good strategy is to start with simple patterns such as compound symmetry or AR(1). Once the general covariance pattern is determined, more complex patterns based on the selected general covariance pattern can be investigated. The more complex pattern would be accepted only if using them leads to a significant improvement in the likelihood.

If there are only a few repeated measurements, there will be little difference between models using different covariance patterns on overall treatment effects. The compound symmetry pattern is likely to be robust and can be used with reasonable confidence.

Example: Covariance pattern models for normal data

We will use the blood pressure data again, but this time consider the repeated measures data. The table below summarizes the number of patients that attended each visit.

Visit	Treatment A	Treatment B	Treatment C	Total
1	106	101	104	311
2	106	100	102	308
3	100	96	94	290
4	95	91	94	280
5	87	88	93	268
6	83	84	91	258

Visualizing correlation patterns graphically

The first graphical tool is visualizing the correlation structure by plotting changes in covariance and correlation among residuals on the same subject at different times over distance between times of observations. The following PROC MIXED program allows us to obtain the correlations and covariances among the residuals.

```
proc mixed data=dbp;  
class trt pat visit;  
model dbp=trt|visit;  
repeated/type=un subject=pat(trt) sscp rcorr;  
ods output covparms=cov;  
ods output rcorr=corr;  
run;
```

The **repeated** statement determines the form of the covariance, **type=un** specifies an unstructured correlation pattern within each pat(trt). **subject=pat(trt)** specifies that errors are correlated within each pat(trt). In other words, the observations on different pat(trt) levels are independent, but observations on the same pat(trt) are not. The **sscp** option causes the $\text{cov}(e_{ijk}, e_{ijk'})$ to be computed directly from the corrected sums of squares and cross products matrix rather than the default REML procedure. The **rcorr** option causes the correlations to be computed. The two **ODS** statements create new SAS data sets containing the covariances and correlations.

The following code sets up the dataset for producing the plot for visualizing correlation patterns

```
data times;  
do time1=1 to 4;  
  do time2=1 to time1;  
    dist=time1-time2;  
    output;  
  end;  
end;  
run;
```

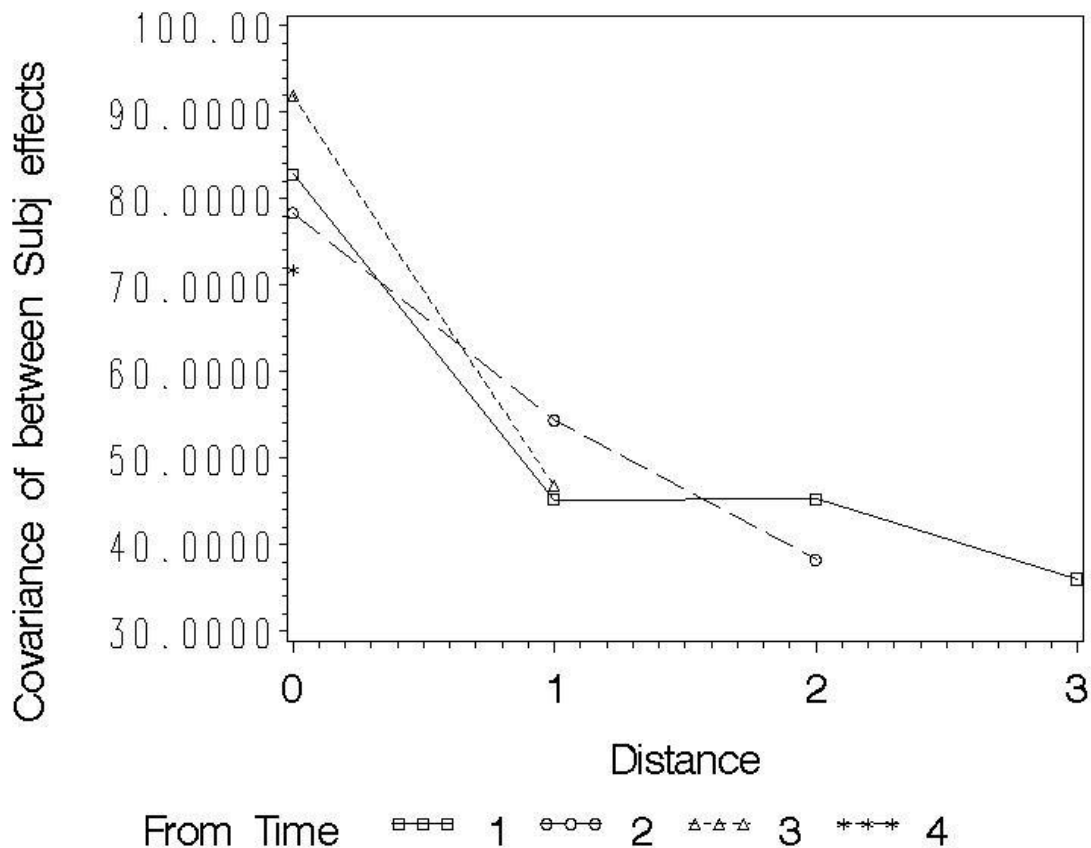
Obs	time1	time2	dist	CovParm	Subject	Estimate
1	1	1	0	UN(1,1)	pat(trt)	82.8090
2	2	1	1	UN(2,1)	pat(trt)	45.1939
3	2	2	0	UN(2,2)	pat(trt)	78.3496
4	3	1	2	UN(3,1)	pat(trt)	45.2571
5	3	2	1	UN(3,2)	pat(trt)	54.3531
6	3	3	0	UN(3,3)	pat(trt)	91.9826
7	4	1	3	UN(4,1)	pat(trt)	36.0270
8	4	2	2	UN(4,2)	pat(trt)	38.2391
9	4	3	1	UN(4,3)	pat(trt)	46.7520
10	4	4	0	UN(4,4)	pat(trt)	71.7496

These are the estimated covariances among residuals on the same subject at different times over distance between times of observations. For example UN(4,1) is the

covariance between the first and fourth time points. The following SAS code plots the above covariances.

```

data covplot;
merge times cov;
proc print;
run;
axis1 value=(font=swiss2 h=2) label=(angle=90 f=swiss h=2 'Covariance of
between Subj effects');
axis2 value=(font=swiss h=2) label=(f=swiss h=2 'Distance');
legend1 value=(font=swiss h=2) label=(f=swiss h=2 'From Time');
symbol1 color=black interpol=join line=1 value=square;
symbol2 color=black interpol=join line=2 value=circle;
symbol3 color=black interpol=join line=20 value=triangle;
symbol4 color=black interpol=join line=4 value=star;
proc gplot data=covplot;
plot estimate*dist=time2/vaxis=axis1 haxis=axis2 legend=legend1;
run;
    
```



The values plotted at distance=0 are the variances among the observations at each of the four visits. These range from about 70 to 90. This small range along with no trend of increasing or decreasing variance with visit, suggests that a covariance pattern with a constant variance over time, such as a Toeplitz or AR(1) will probably be adequate. As

the distance between pairs of observations increases, the covariance tends to decrease, again suggesting a Toeplitz or AR(1) covariance pattern. Even though the plot narrows the choice of which covariance patterns may be appropriate, we will fit and compare six different covariance patterns to our data.

Model	Covariance pattern
1	compound symmetry
2	first order autoregressive (AR(1))
3	Toeplitz
4	General
5	Separate compound symmetry for each treatment group
6	Separate Toeplitz pattern for each treatment group

Treatment, time, treatment*time and baseline effects will be included as fixed effects. The following SAS code fits the compound symmetry model.

```
proc mixed data=dbp;
class trt pat visit;
model dbp=trt|visit dbp0;
repeated visit/type=cs subject=pat r rcorr;
run;
```

The following output is produced.

Model Information			
Data Set		WORK.DBP	
Dependent Variable		dbp	
Covariance Structure		Compound Symmetry	
Subject Effect		pat	
Estimation Method		REML	
Residual Variance Method		Profile	
Fixed Effects SE Method		Model-Based	
Degrees of Freedom Method		Between-Within	
Dimensions			
Covariance Parameters		2	
Columns in X		21	
Columns in Z		0	
Subjects		288	
Max Obs Per Subject		4	
Number of Observations			
Number of Observations Read		1092	
Number of Observations Used		1092	
Number of Observations Not Used		0	
Iteration History			
Iteration	Evaluations	-2 Res Log Like	Criterion

```

0          1      7794.69107620
1          2      7463.35292580      0.00000049
2          1      7463.35156497      0.00000000
    
```

Convergence criteria met.

Estimated R Matrix for pat 1

```

1      76.3680      40.1639      40.1639      40.1639
2      40.1639      76.3680      40.1639      40.1639
3      40.1639      40.1639      76.3680      40.1639
4      40.1639      40.1639      40.1639      76.3680
    
```

Estimated R Correlation Matrix for pat 1

```

1      1.0000      0.5259      0.5259      0.5259
2      0.5259      1.0000      0.5259      0.5259
3      0.5259      0.5259      1.0000      0.5259
4      0.5259      0.5259      0.5259      1.0000
    
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
CS	pat	40.1639
Residual		36.2041

Fit Statistics

-2 Res Log Likelihood	7463.4
AIC (smaller is better)	7467.4
AICC (smaller is better)	7467.4
BIC (smaller is better)	7474.7

The next three models use the same SAS code, except the type= is modified as follows:

type=AR(1) for AR(1)

type=toep for Toeplitz

type=un for general or unstructured

The following are the covariance parameters and information criterion

AR(1)

```

Residual      76.1890
Estimated R Correlation Matrix for AR(1)

1.0000      0.5728      0.3281      0.1879
0.5728      1.0000      0.5728      0.3281
0.3281      0.5728      1.0000      0.5728
0.1879      0.3281      0.5728      1.0000
    
```

-2 Res Log Likelihood	7485.3
AIC (smaller is better)	7489.3
AICC (smaller is better)	7489.3
BIC (smaller is better)	7496.6

Toeplitz

```

Residual      76.3091
Estimated R Correlation Matrix for Toeplitz

1.0000      0.5722      0.4818      0.4620
0.5722      1.0000      0.5722      0.4818
0.4818      0.5722      1.0000      0.5722
    
```

0.4620 0.4818 0.5722 1.0000
 -2 Res Log Likelihood 7450.6
 AIC (smaller is better) 7458.6
 AICC (smaller is better) 7458.6
 BIC (smaller is better) 7473.2

General or Unstructured

UN(1,1) pat 75.8677
 UN(2,2) pat 70.9100
 UN(3,3) pat 86.2167
 UN(4,4) pat 73.3582

Estimated R Correlation Matrix for General or Unstructured

1.0000	0.5154	0.4781	0.4554
0.5154	1.0000	0.6108	0.4986
0.4781	0.6108	1.0000	0.6066
0.4554	0.4986	0.6066	1.0000

-2 Res Log Likelihood 7442.3
 AIC (smaller is better) 7462.3
 AICC (smaller is better) 7462.5
 BIC (smaller is better) 7498.9

The following table summarizes the fit information for the four models

Model	-2 Res Log (L)	AIC	BIC
1 compound symmetry	7463.4	7467.4	7474.7
2 first order autoregressive (AR(1))	7485.3	7489.3	7496.6
3 Toeplitz	7450.6	7458.6	7473.2
4 General	7442.3	7462.3	7498.9

For all of the models, correlations between visits are positive, indicating that it is important to take into account the correlations between repeated measures. Comparing the fit statistics among these first four covariance patterns, it appears that the compound symmetry (the simplest model with 2 covariance parameters) and the Toeplitz (with 4 covariance parameters) covariance patterns fit the data better than either the AR(1) or general covariance patterns. The compound symmetry model is nested within the Toeplitz pattern, so they can be compared statistically using the likelihood ratio test. The likelihood ratio test statistic would be

$$2(\log(L_1)-\log(L_2)) \sim \chi_{df}^2 = 2(.5*7463.4-.5*7450.6)=7463.4-7450.6=12.8 = \chi_2^2.$$

The p-value associated with this $\chi_2^2 = .002$, indicating that the Toeplitz structure is a significant improvement over the simpler compound symmetry structure.

Model 5's covariance pattern allows for differences in variances and covariances among the treatments. The SAS code for fitting this model with a compound symmetry covariance pattern follows

```
proc mixed noclprint data=dbp;
class trt pat visit;
```

```
model dbp=trt|visit dbp0/ddfm=satterth;  
repeated visit/type=cs subject=pat group=trt r=1,3,4 rcorr=1,3,4;  
run;
```

There are a couple of additions/changes in the repeated statement. **group=trt** allows for estimation of the parameters by treatment. **r=1,3,4 rcorr=1,3,4** indicates that the covariance matrices and correlation matrices for Patients 1(Trt C), 3 (Trt B), and 4 (Trt A) be printed out. Following are the results for Model 5

Estimated R Matrix for pat 1 - Treatment C

1	76.1484	47.9904	47.9904	47.9904
2	47.9904	76.1484	47.9904	47.9904
3	47.9904	47.9904	76.1484	47.9904
4	47.9904	47.9904	47.9904	76.1484

Estimated R Correlation Matrix for pat 1 - Treatment C

1	1.0000	0.6302	0.6302	0.6302
2	0.6302	1.0000	0.6302	0.6302
3	0.6302	0.6302	1.0000	0.6302
4	0.6302	0.6302	0.6302	1.0000

Estimated R Matrix for pat 3 - Treatment B

1	68.1970	26.7774	26.7774	26.7774
2	26.7774	68.1970	26.7774	26.7774
3	26.7774	26.7774	68.1970	26.7774
4	26.7774	26.7774	26.7774	68.1970

Estimated R Correlation Matrix for pat 3 - Treatment B

1	1.0000	0.3926	0.3926	0.3926
2	0.3926	1.0000	0.3926	0.3926
3	0.3926	0.3926	1.0000	0.3926
4	0.3926	0.3926	0.3926	1.0000

Estimated R Matrix for pat 4 - Treatment A

1	85.0507	45.8115	45.8115	45.8115
2	45.8115	85.0507	45.8115	45.8115
3	45.8115	45.8115	85.0507	45.8115
4	45.8115	45.8115	45.8115	85.0507

Estimated R Correlation Matrix for pat 4 - Treatment A

1	1.0000	0.5386	0.5386	0.5386
2	0.5386	1.0000	0.5386	0.5386
3	0.5386	0.5386	1.0000	0.5386
4	0.5386	0.5386	0.5386	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate
Variance	pat	trt A	39.2393
CS	pat	trt A	45.8115
Variance	pat	trt B	41.4195
CS	pat	trt B	26.7774
Variance	pat	trt C	28.1580
CS	pat	trt C	47.9904

Fit Statistics

-2 Res Log Likelihood	7447.5
AIC (smaller is better)	7459.5
AICC (smaller is better)	7459.6
BIC (smaller is better)	7481.5

Looking at the results, the pattern for compound symmetry differ among the treatments, indicating that separate covariances for each treatment group may be necessary. Because Model 3 (Toeplitz) fit significantly better than Model 1 (compound symmetry), a Toeplitz model fitting for separate covariances for each treatment group should be fitted next (Model 6) with the following SAS code and results.

```
proc mixed data=dbp;
class trt pat visit;
model dbp=trt|visit dbp0;
repeated visit/type=toep subject=pat group=trt r=1,3,4 rcorr=1,3,4;
lsmeans trt/ diff pdiff cl;
```

Estimated R Matrix for pat 1 - Treatment C

1	76.1169	52.7624	46.4925	35.3745
2	52.7624	76.1169	52.7624	46.4925
3	46.4925	52.7624	76.1169	52.7624
4	35.3745	46.4925	52.7624	76.1169

Estimated R Correlation Matrix for pat 1

1	1.0000	0.6932	0.6108	0.4647
2	0.6932	1.0000	0.6932	0.6108

3	0.6108	0.6932	1.0000	0.6932
4	0.4647	0.6108	0.6932	1.0000

Estimated R Matrix for pat 3 - Treatment B

1	68.2100	28.9323	22.4760	28.6921
2	28.9323	68.2100	28.9323	22.4760
3	22.4760	28.9323	68.2100	28.9323
4	28.6921	22.4760	28.9323	68.2100

Estimated R Correlation Matrix for pat 3

1	1.0000	0.4242	0.3295	0.4206
2	0.4242	1.0000	0.4242	0.3295
3	0.3295	0.4242	1.0000	0.4242
4	0.4206	0.3295	0.4242	1.0000

Estimated R Matrix for pat 4

1	84.9809	49.5186	41.0737	42.2137
2	49.5186	84.9809	49.5186	41.0737
3	41.0737	49.5186	84.9809	49.5186
4	42.2137	41.0737	49.5186	84.9809

Estimated R Correlation Matrix for pat 4

1	1.0000	0.5827	0.4833	0.4967
2	0.5827	1.0000	0.5827	0.4833
3	0.4833	0.5827	1.0000	0.5827
4	0.4967	0.4833	0.5827	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate
Variance	pat	trt A	84.9809
TOEP(2)	pat	trt A	49.5186
TOEP(3)	pat	trt A	41.0737
TOEP(4)	pat	trt A	42.2137
Variance	pat	trt B	68.2100
TOEP(2)	pat	trt B	28.9323
TOEP(3)	pat	trt B	22.4760
TOEP(4)	pat	trt B	28.6921
Variance	pat	trt C	76.1169
TOEP(2)	pat	trt C	52.7624
TOEP(3)	pat	trt C	46.4925
TOEP(4)	pat	trt C	35.3745

Fit Statistics

-2 Res Log Likelihood	7424.0
AIC (smaller is better)	7448.0
AICC (smaller is better)	7448.3
BIC (smaller is better)	7491.9

Both the simple Toeplitz model (Model 3), and the compound symmetry with treatment model (Model 5) are nested within Model 6, so we can test to see if Model 6 fits significantly better than the other two models.

Model 3 (4 parameters) vs Model 6 (12 parameters)

$$2(\log(L_1) - \log(L_2)) \sim \chi_{df}^2 = 7450.6 - 7424.0 = 26.6 = \chi_8^2 \quad (p=0.0008)$$

Model 5 (6 parameters) vs Model 6 (12 parameters)

$$2(\log(L_1)-\log(L_2)) \sim \chi^2_{df} = 7447.5-7424.0=23.5= \chi^2_6 \quad (p=0.0006)$$

Therefore, we are statistically justified in using the more complex covariance pattern. This may partly be due to the large size of the trial, which allows us to estimate the covariances with reasonable accuracy.