Random 3: Repeated Measures

Outline:

1. One-factor repeated measures – Basic Analysis
2. One-factor repeated measures – Covariance Structures
3. Two-factor repeated measures
4. Using baseline (pre-treatment) as a covariate

Examples:

1. Exercise: One-factor repeated measures
2. Cholesterol1: One-factor repeated measures
3. Cholesterol2: Two-factor repeated measures
4. Cats: Repeated measures with Covariate

0. Some comments about lme4 vs nlme

In this group of notes, we will use lme() from the nlme package to fit some mixed models. Both lme4 and nlme were developed by Doug Bates (as well as others!) but use different syntax.

Both packages: can be used to fit basic mixed models.

lme4 benefits:

- Type3 tests with Kenward Rogers df using lmerTest
- Generalized linear mixed models (ex: logistic regression with random effects)

nlme benefits:

- Correlation structures
- Modeling of unequal variances.

1. One Factor Repeated Measures- Basic Analysis

The term "**repeated measures**" design refers to any design in which an experimental unit (EU) is measured more than once. In such designs, multiple measurements on the same EU will not be independent. They will be correlated.

Randomized complete block designs also can be thought of as repeated measures designs when the blocks are assumed randomly selected. Multiple measurements on the same block are correlated.

Much of what is described in books as analysis of repeated measures designs involves nothing beyond what we have already studied.

We will start with a one factor repeated measures analysis analyzed as an RCB.

Exercise Example: One-Factor Repeated Measures (O&L Example 18.2):

An exercise physiologist designed a study to evaluate the impact of the steepness of running courses on the peak heart rate (Y=PHR). There are **4 courses**: flat, slightly steep, moderately steep and very steep. A total of **20 runners** will run each of the courses in a randomly assigned order. There will be sufficient time between runs so that there should not be any carryover effect. The (basic) analysis is the same as RCB.

Let Y_{ij} be the response for the ith course for the jth runner:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

$$\beta_j \sim N(0, \sigma^2_\beta),\ \varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$$

Repeated Measures Terminology

• The **subject** is an experimental unit.
• When a variable is measured repeatedly for all subjects across a set of conditions, this set of conditions is called a **within-subjects factor**.
• When a variable is measured on independent groups of subjects, where each group is exposed to a different condition, the set of conditions is called a **between-subjects factor.**
• Note that when you change levels of a "**between subjects**" factor, you must be talking about a different subject. When you change levels of a "**within subjects**" factor, you could be talking about the same subject.

For the Exercise Data:

Subject = Runner

Within-Subjects Factor = Course (4 levels)

Between-Subjects Factor = N/A

Our primary interest is in comparing the mean response for the 4 courses.

We will run the analysis 2 different ways (for illustration), but the results will be exactly the same:

1. lmer (from lme4):

```
lmer(Y ~ Course + (1|Runner)
```

2. lme (from nlme):

```
lme(Y ~ Course, random = ~1|Runner)
```

Compound Symmetry:

We now look at the implications of the assumptions of the RCB model by looking at the covariances of measurements on the same subject:

$$\begin{aligned} Var(y_{ij}) &= Var(\mu + \alpha_i + \beta_j + \varepsilon_{ij}) \\ &= Var(\beta_j) + Var(\varepsilon_{ij}) \\ &= \sigma_\beta^2 + \sigma_\varepsilon^2 \end{aligned}$$

$$\begin{aligned} Cov(y_{ij}, y_{i'j}) &= Cov(\mu + \alpha_i + \beta_j + \varepsilon_{ij}, \mu + \alpha_{i'} + \beta_j + \varepsilon_{i'j}) \\ &= Cov(\beta_j, \beta_j) \\ &= Var(\beta_j) \\ &= \sigma_\beta^2 \end{aligned}$$

Compound Symmetry Continued:

Hint about "Covariance": Think of a covariance as a correlation, without removing the scale:

$$Corr(y_{ij}, y_{i'j}) = \frac{Cov}{(Std.dev)(Std.dev.)} = \frac{\sigma_\beta^2}{\sqrt{\sigma_\beta^2 + \sigma_\varepsilon^2}\sqrt{\sigma_\beta^2 + \sigma_\varepsilon^2}} = \frac{\sigma_\beta^2}{\sigma_\beta^2 + \sigma_\varepsilon^2} = \rho$$

The above correlation (ρ) is called the "intra-subject" correlation; that is, the correlation between two measurements on the same subject. (When fit using a random subject effect this correlation cannot be negative.)

The covariance of measurements on different subjects is always zero:

$$Cov(y_{ij}, y_{i'j'}) = 0$$

Compound Symmetry Continued:

Compound symmetry is implied by the Basic/RCB type analysis. The compound symmetry assumption implies that there is constant correlation between observations on the same subject.

In some cases this is reasonable and the RCB type analysis is appropriate. In other cases (for example measurements across time), it may not be reasonable.

We will discuss how to check this assumption and how to allow different assumptions in the next section.

Sphericity is a less restrictive assumption than compound symmetry.

Friedman's Test: A non-parametric alternative to the RCB analysis

Friedman's test is similar to the Kruskal-Wallis test, but ranking is done within each subject (or block).

In R, Friedman's test can be done using `friedman.test()` from the stats package

2. One-factor repeated measures – Covariance Structures

Cholesterol1 Example: One factor Repeated Measures

Twelve subjects were given a drug designed to lower serum cholesterol. Their cholesterol was measured at the end of six four-week periods. (Period 1 is baseline, before treatment).

Note: Unlike the exercise data, the periods come in time order and cannot be randomized. We use the same model.

Let y_{ij} = cholesterol for the i^{th} period $i = 1, \ldots, a = 6$

j^{th} subject $j = 1, \ldots, r = 12$

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

μ and α_i are fixed effects

β_j are $N(0, \sigma^2_\beta)$ ε_{ij} are $N(0, \sigma^2)$

β_j, ε_{ij} are all independent.

Cholesterol Example:

Subject = Subject

Within subjects factors = Period

Between subjects factors = N/A

Analysis plan:

We will first analyze the data using the basic model (RCB), which assumes compound symmetry. But we will now also consider alternative "correlation structures".

To see the differences between the assumptions, we will write out the covariance matrix for a single subject. In R, the `getVarCov()` function will do this.

Covariance matrix for all six periods on the same subject with Compound Symmetry (correlation = corCompSymm):

$$Cov\begin{pmatrix} y_{1j} \\ y_{2j} \\ y_{3j} \\ y_{4j} \\ y_{5j} \\ y_{6j} \end{pmatrix} = \begin{bmatrix} \sigma_\beta^2+\sigma^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 \\ \sigma_\beta^2 & \sigma_\beta^2+\sigma^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 \\ \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2+\sigma^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 \\ \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2+\sigma^2 & \sigma_\beta^2 & \sigma_\beta^2 \\ \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2+\sigma^2 & \sigma_\beta^2 \\ \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2 & \sigma_\beta^2+\sigma^2 \end{bmatrix}$$

<u>Assumptions implied by this model:</u>

1. The diagonal elements are the same. ("homogeneity of variance")
2. All off-diagonals are the same.

How reasonable is the "compound symmetry" (CS) model?

This model assumes that the covariance between the measurements at periods 1 and 2 is the same as the covariance between periods 1 and 3. That's what makes the standard errors for comparing any two periods the same in the output.

Does this seem reasonable? Wouldn't you expect measurements closer in time to be more highly correlated? How about the correlation between periods 1 and 6? Periods 1 and 10?

Since the assumption of "compound symmetry" seems questionable, we will consider other possibilities.

The corClasses help lists about 10 correlation structures. (SAS Proc Mixed offers >25.) Here we consider just a few of the alternative models.

Alternative Assumption 1: **Unstructured covariance (correlation = corSymm)**

In this model there are no assumptions about the covariance matrix, except that Cov(X,Y)=Cov(Y,X), for any X and Y.

$$Cov\begin{pmatrix} y_{1j} \\ y_{2j} \\ y_{3j} \\ y_{4j} \\ y_{5j} \\ y_{6j} \end{pmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} & \sigma_{16} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} & \sigma_{25} & \sigma_{26} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} & \sigma_{35} & \sigma_{36} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 & \sigma_{45} & \sigma_{46} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_5^2 & \sigma_{56} \\ \sigma_{61} & \sigma_{62} & \sigma_{63} & \sigma_{64} & \sigma_{65} & \sigma_6^2 \end{bmatrix}$$

Where: $\sigma_{ij} = \sigma_{ji} = Cov(y_{1j}, y_{2j})$

Notes:

1. As shown on the previous slide, this model does not assume homogeneity of variance. The variance for each period could be potentially different. However, in the current R example the variances are forced to be the same. This can be changed using the varIndent option, but not shown.

2. This model does not assume "compound symmetry": the covariance between any two periods is potentially different.

3. Our example has six periods; there are 17 covariance parameters. The number of parameters to estimate increases rapidly with the number of periods.

4. Using corSymm is a good way to see what structure might be reasonable. My opinion: I typically do NOT want to choose this for my final model because of the large number of parameters!

Alternative Assumption 2: **Autoregressive Covariance (correlation = corAR1)**

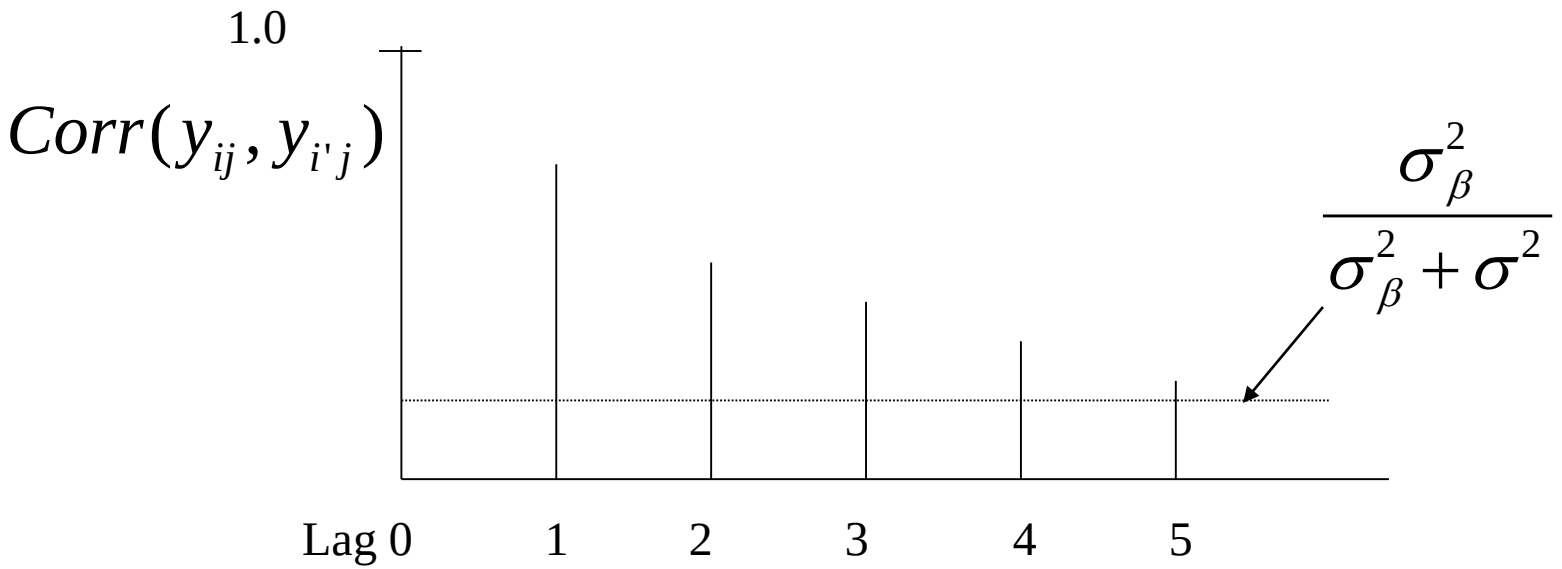
Assumes that the covariance between two periods decreases exponentially, depending on the time between periods.

Note: $0 \leq \rho \leq 1$

$$\begin{bmatrix}
\sigma_\beta^2+\sigma^2 & \sigma_\beta^2+\rho\sigma^2 & \sigma_\beta^2+\rho^2\sigma^2 & \sigma_\beta^2+\rho^3\sigma^2 & \sigma_\beta^2+\rho^4\sigma^2 & \sigma_\beta^2+\rho^5\sigma^2 \\
\sigma_\beta^2+\rho\sigma^2 & \sigma_\beta^2+\sigma^2 & \sigma_\beta^2+\rho\sigma^2 & \sigma_\beta^2+\rho^2\sigma^2 & \sigma_\beta^2+\rho^3\sigma^2 & \sigma_\beta^2+\rho^4\sigma^2 \\
\sigma_\beta^2+\rho^2\sigma^2 & \sigma_\beta^2+\rho\sigma^2 & \sigma_\beta^2+\sigma^2 & \sigma_\beta^2+\rho\sigma^2 & \sigma_\beta^2+\rho^2\sigma^2 & \sigma_\beta^2+\rho^3\sigma^2 \\
\sigma_\beta^2+\rho^3\sigma^2 & \sigma_\beta^2+\rho^2\sigma^2 & \sigma_\beta^2+\rho\sigma^2 & \sigma_\beta^2+\sigma^2 & \sigma_\beta^2+\rho\sigma^2 & \sigma_\beta^2+\rho^2\sigma^2 \\
\sigma_\beta^2+\rho^4\sigma^2 & \sigma_\beta^2+\rho^3\sigma^2 & \sigma_\beta^2+\rho^2\sigma^2 & \sigma_\beta^2+\rho\sigma^2 & \sigma_\beta^2+\sigma^2 & \sigma_\beta^2+\rho\sigma^2 \\
\sigma_\beta^2+\rho^5\sigma^2 & \sigma_\beta^2+\rho^4\sigma^2 & \sigma_\beta^2+\rho^3\sigma^2 & \sigma_\beta^2+\rho^2\sigma^2 & \sigma_\beta^2+\rho\sigma^2 & \sigma_\beta^2+\sigma^2
\end{bmatrix}$$

This model has three covariance parameters, no matter how many periods. This is only a slight increase in number of parameters over the "compound symmetry" model, which has two parameters.

The relationship between correlation and lag (distance between periods) looks like this:



1. This model <u>does</u> assume homogeneity of variance.
2. The correlation between two measurements on the same subject decreases exponentially with the number of time periods separating the two measurements.

Comparison of treatment means:

Using the Basic (CS) model, the standard errors for comparing period means were the same (8.78), regardless of the periods being compared.

Using the AR(1) model, the standard errors for comparing the period means now depend on the number of periods apart the means are:

Adjacent periods: SE(diff)=6.424
Two periods apart: SE(diff)=8.628
Five periods apart: SE(diff)=11.822

Note: When the design is balanced the estimated emmeans are not changed by the assumption of variance model, but the standard errors are greatly affected by the assumptions about the variance model.

Comments about the lme() correlation option:

1. If a study has just two time points (pre, post?) then there is no need to consider different covariance structures. All options (that we considered) reduce to the same model!
2. **For the corAR1:** Watch out for the ordering of the time levels! Check the factor level information. Here the ordering is obvious: Periods 1, 2, 3, 4, 5, 6. But in some cases labeling can cause problems (Ex: Day0, Week4, Month2).
3. corAR1 models assume equal spacing of periods. If the periods are not equally spaced other approaches are available (for example: corCAR1).
3. Several options are available for modeling spatial correlation: corExp, corGaus, corLin, etc.

Plan for the Analysis:

1. Examine the estimates of the covariances for the unstructured model (corSymm) and look for patterns.
2. Choose a covariance structure using:

   A. AIC values

   B. When one model can be described as a restricted version of another model, compare that pair of models using the "Restricted Likelihood Ratio Test".
3. Address research questions by using the selected model to estimate and compare means (corresponding to fixed effects).

IMPORTANT NOTES ABOUT AIC CALCULATIONS:

1. Recall that AIC = -2loglik + 2p where p is the # of parameters.

2. For these comparisons (because REML is used for estimation), the models being compared need to have the same fixed effects.

3. The count of parameters includes both fixed effects and variance components.

Model comparison using AIC:

Model	-2LogLik	Fp	Rp	p	AIC
corCompSymm	644.0632	6	2	8	660.0632
corSymm	606.5514	6	17	23	652.5514
corAR1	626.5804	6	3	9	644.5804

corAR1 has the lowest AIC.

Restricted Likelihood Ratio Test:

Two models can be compared using a procedure that is similar to the F-test we used for multiple regression and ANOVA:

1. A larger model (the "full" model") can be compared to a smaller model (the "restricted" model) by comparing the difference between the -2 log likelihood values to a chi-square distribution. df=difference in number of parameters.

   H0: The restricted (reduced) model fits (compared to the full model).

   HA: The restricted (reduced) model does NOT fit (compared to the full model).

2. When using REML estimation, both models must have the same set of fixed effects.

Some Likelihood Ratio Tests for the Cholesterol Data:

Comparison #1: CompSymm vs Symm
Chi-square = 644.1 – 606.6 = 37.5
df = 17 – 2 = 15, p-value = 0.001
Conclusion: Reject H0; CompSymm model is not adequate.

Comparison #2: AR1 vs Symm
Chi-square = 626.6 – 606.6 = 20
df = 17 – 3 = 14, p-value = 0.130
Conclusion: Fail to Reject H0; AR(1) model is adequate.

Comparison #3 CompSymm vs AR1
Chi-square = 644.1 – 626.6 = 17.5
df = 3 - 2 = 1, p-value < 0.001
Conclusion: Reject H0; CompSymm model is not adequate.

1. Any of these model can be thought of as a restricted version of the unrestricted model (corSymm).
2. The AIC may be used to compare any two of these models (whether one model is a restricted version of the other, or not).
3. Reminder: AIC may be used to compare REML models, as long as the models being compared have the same fixed effects.
4. AIC = -2 Log Res Lik +2p. A restricted model with 1 fewer parameters will have lower AIC whenever the Lik Ratio Test statistic exceeds 2 ($p<0.157$). Models with 2, 3, or 4 fewer parameters will have lower AIC whenever the Lik Ratio Test statistic exceeds 4 ($p<0.135$), 6 ($p<0.111$), and 8 ($p<0.092$), respectively. Therefore "lowest AIC" is a less strict criterion for adding parameters than the usual hypothesis test.

3. Two Factor Repeated Measures Design

The previous cholesterol example was only part of the study of the cholesterol drug. Additionally, a control group with thirteen subjects was given a placebo and measured over the same six periods. (Period 1 is baseline: before treatment).

Let y_{ijk} = cholesterol

for the i^{th} trt, i=1,...,a=2

j^{th} subject, j=1,...,r_i =12 or 13

k^{th} period k=1,...,b=6

$$y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \tau_k + (\alpha\tau)_{ik} + \varepsilon_{ijk}$$

μ , α_i, τ_k and $(\alpha\tau)_{ik}$ are fixed effects

$\beta_{j(i)}$ are $N(0,\sigma^2_{\beta(\alpha)})$ ε_{ijk} are $N(0,\sigma^2)$

$\beta_{j(i)}$, ε_{ijk} are all independent.

Cholesterol Example (now w/Trts):

Subject = Subject (nested within Trt)

Within subjects factors = Period

Between subjects factors = Trt

Remember: When you change levels of a "**between subjects**" factor, you must be talking about a different subject. When you change levels of a "**within subjects**" factor, you could be talking about the same subject.

Analysis Plan: We will again consider a number of covariance structures. We compare AIC values for the models. Then estimate and compare means (corresponding to fixed effects).

Relationship to Split-Plot Design:

This model looks like a split-plot design (without blocks), where a subject is a whole plot, and a period is a subplot. It is not really a split-plot design, because periods occur in a specific order and cannot be randomly assigned.

Remember that we get smaller SE's for comparisons of the sub-plot factor (within-subjects factor = Period) and larger SE's for comparisons of the whole-plot factor (between-subjects factor = Trt). (This is when using basic CS, not necessarily other methods.)

Split plot:		Repeated Measures
whole plot	=	subject
whole plot factor	=	between subjects factor
subplot factor	=	within subjects factor
subplot	=	1 measurement on a subject

Model comparison using AIC:

Model	-2LogLik	Fp	Rp	p	AIC
CS	1215.759	12	2	14	1243.759
corSymm	1184.5518	12	17	29	1242.5518
corAR1	1203.6148	12	3	15	1233.6148

corAR1 has the smallest AIC.

Conclusions about this example:

1. The basic CS model is inadequate, because measurements at periods that are close together in time are correlated more highly than measurements at periods that are farther apart in time.

2. The corAR1 model seems to adequately explain the correlation of time periods.

3. There is no evidence of a difference between the active drug and placebo groups, although there is some evidence of a period effect, with later periods lower than baseline. (White coat effect?)

4. Other comparisons that might be of interest:

   A. Compare across periods for a single Trt (placebo or test).

   ```
   emmeans(Model, dunnett ~ period|trt)
   ```

   Note: I have chosen to use Dunnett adjustment to compare each later periods to baseline.

   B. Look for differences between Trts at each period.

   ```
   emmeans(Model, pairwise ~ trt|period)
   ```

5. Comparing the CS and AR(1) models: Comparison (A) is effected by the choice of model, but Comparison (B) is relatively unaffected by the choice of model.

6. Using the either model, the SE for comparing Trts is larger than the SE for comparing Periods. This is because Period is the within-subject (or sub-plot) factor and Trt is the between-subjects (or whole plot) factor.

4. Including a pre-treatment time point (baseline) as a covariate

In studies with repeated measures, it is common to have a pre-treatment measure called a baseline. This permits using each subject as it "own control" to asses the effect of treatment over time.

We can include the baseline value as a covariate in the repeated measures analysis.

This is similar to (but not the same as) taking the difference between each later time point vs baseline.

Example: A study was done to compare 2 Treatments (A and C) for diabetes in cats. 6 Cats were randomly assigned to each treatment. Blood sugar (Y) measurements were taken at 4 equally spaced Times (0, 1, 2, 3). The first time point (Time=0) was pre-treatment and serves as a baseline.

Subject = Cat (nested within Trt)

Within subjects factors = Time

Between subjects factors = Trt

We will try 3 approaches:

1. Basic Repeated measures
2. Repeated measures with Time0 (baseline) as covariate
3. Repeated measures on Difference values

*NOTE: Different covariance structures were NOT investigated, but certainly could be.

Two-factor Repeated Measures Analysis with Covariate

Let y_{ijk} represent the blood sugar reading for Trt i (i=A or C), subject j (1 through 6) and period k (1-3). Let x_{ij} represent the baseline (time 0= pre-treatment) response for subject j from group i.

$$y_{ijk} = \mu + \gamma x_{ij} + \alpha_i + \beta_{j(i)} + \tau_k + (\alpha\tau)_{ik} + \varepsilon_{ijk}$$

Comments on the Cats Example:

1. Covariates in the repeated measures analysis can take forms other than a baseline value. For example initial weight or age.
2. The baseline covariate (or difference) models "adjust" for the baseline values of each individual. This approach can be used to account for groups that showed initial differences.
3. One result of using the baseline covariate or difference models is that we no longer get direct comparisons between the groups as baseline. If needed, this can be done using a two-sample t-test (for example).

4. In this example, the baseline covariate and difference models give very similar conclusions. But the interpretation of results from the two models is different. Sometimes interpretation is easier using one model or the other.
5. Note that in the "basic" repeated measures analysis, the Cat variance was estimated to be 50.18. In the baseline covariate model, the Cat variance was estimated to be 4.59. In the difference model, the Cat variance was estimated to be 7.41. This drop in the estimated variance is not surprising because the baseline value for each cat is accounting for a lot of the cat to cat variability. Sometimes the estimated variance even goes to zero.

Example write-up for Cats Analysis (with covariate):

Analysis was done using R and the lme4, lmerTest and emmeans packages (provide references!). A mixed model was fit using Y as the response. Fixed effects included treatment (A or C), time (1, 2, 3) plus treatment*time interaction. Baseline reading was included as a covariate. Subject was included as a random effect to account for repeated measures.