Longitudinal models and missing data

See the course notes (*Longitudinal models and missing data*) for more detail. This is new stuff! ☺

1 Introduction

- There is a wealth of research devoted to missing data, its impact on results, and how to deal with it.
- Missing data are particularly a problem for longitudinal studies because often subjects start in the study, but for one reason or another they dropout.
- In this chapter, we first discuss how mixed models can naturally account for missing data by taking correlation between responses into account. We then discuss types of data that make it harder or easier to account for missing data, and define general types of mechanisms for missingness commonly discussed in the literature.

2 How missing data is handled in multivariate versus univariate procedures

See course notes.

3 Mixed models and estimation in light of missing data

3.1 Relationship between missing data and correlation

Consider outcome data (*Y*) collected on subjects at 2 time points (Visit 1 and 2). Three of the subjects have missing data for Visit 2. Figure 1 shows a plot of the data.

Figure 1

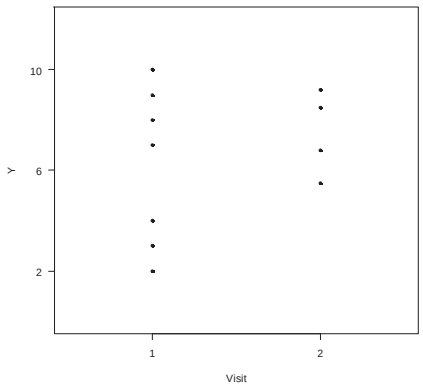
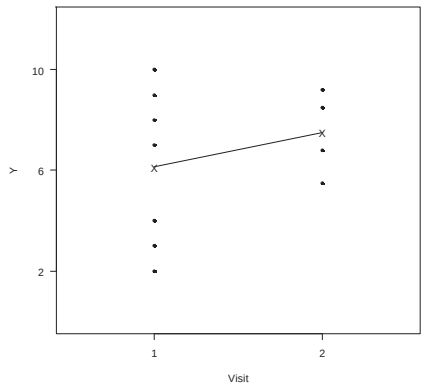


Figure 2



If the repeated measures are ignored and all data used, then there appears to be an increase over time. These would be the estimates obtained using a linear model if no correlation for repeated measures is taken into account in the model (Figure 2).

Code to get estimates for Figure 2 when correlation is ignored:

```
proc mixed data=test;
class id visit;
model y= visit / solution;
lsmeans visit; run;
```

```
Least Squares Means
Effect visit  Estimate    SE      DF t Value Pr>|t|
Visit    1    6.1429    1.0331   9  5.95    0.0002
Visit    2    7.5000    1.3666   9  5.49    0.0004
```

- Now consider identifying the repeated measures within subjects (below). The dashed lines in Figure 3 show the true progression for the subjects with missing data; I have included the missing values with open circles, although the analyst does not observe them.

Figure 3

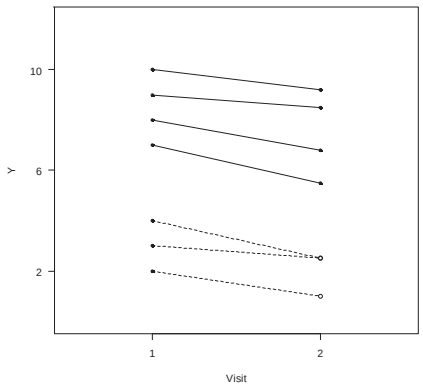
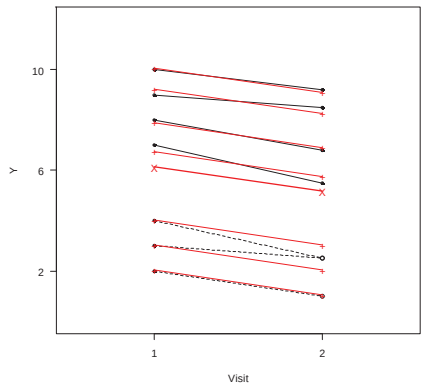


Figure 4



- If we include a random intercept for subjects in the model, we get subject estimates that make sense. In Figure 4, the red lines are the predicted values from the mixed model overlaid on the actual data (black lines). Note that we get predicted values even for the missing y values, although they were not used or imputed for the model fit. The thicker red line with X's at the 2 visits is the population average estimated slope (see SAS code and associated output below for these estimates). Note that it decreases, which is consistent with the actual data when the missing values are taken into account.

- These data illustrate the importance of identifying subjects in the analysis and taking correlation within subjects into account. This particular model works in the situation that the slopes are generally consistent between subjects and not dependent on the starting value. This illustrates that we can get accurate estimates of subject values as well as the population-average fit in the presence of missing data if the model is appropriate for the data.

- Note that we would get the same population-average fit if we included a REPEATED statement with the CS covariance structure defined instead of the random intercept, however we would then not be able to get subject-specific estimates. Note that the standard errors on the mean estimates for the two time points are similar, around 1.2, whereas in the previous model fit that assumed independent data, the SE was larger at Visit 1 than Visit 2; this is a result of our assumed model that takes repeated measures into account.


```
proc mixed data=test;
class id visit;
model y= visit / solution
outp=preddy;
random intercept /
subject=id solution;
lsmeans visit;
estimate 'change over time'
visit -1 1; run;
```

```
Least Squares Means

Effect visit Estimate  SE   t Value   Pr>|t|
visit    1   6.1429  1.1984  35.13    0.0144
visit    2   5.1656  1.2070  34.28    0.0234
```

3.2 Systematic differences that are informative or not

- If the subjects that had missing data for Visit 2 were systematically different than those with complete data, then there is really no way we get good estimates unless the observed or other data informs us about the systematic differences. Below are examples of this (Figures 5 and 6). Figure 5 shows that subjects who miss Visit 2 have systematic differences from the ‘completers’; but the analyst does not observe this.

- Without other information, they will not be able to accurately estimate progression for these subjects, and considering the average progression for all 7 subjects, we will have bias in our estimate when only using the top 4.

- Such data are informative since not knowing them will affect our estimation. [More formally, such data are likely to be *missing not at random* (MNAR).]

- In particular, the average change will be lower if only the top 4 subjects are used in estimation, and would be increased if the analyst had observed all 7 values. Even if subjects with missing values had a mean Visit 1 measure that was similar to that of the 'completers', we could not accurately estimate the slope of the missing subjects or the combined slope (if relevant) due to the systematic difference in slopes between completers and dropouts.

- Now consider data in Figure 6, where the slope change is related to the Visit 1 value; in this case it is possible that we can obtain accurate estimates by employing the (known) relationship between Visit 1 and 2 values.

Figure 5

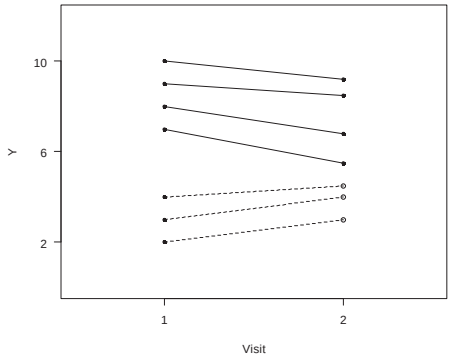
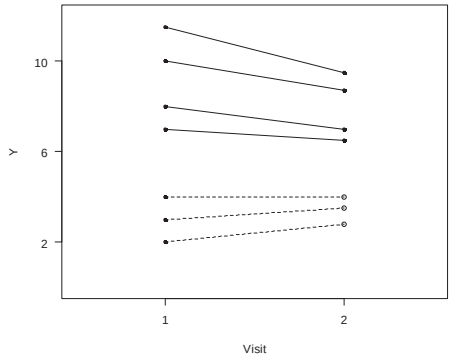


Figure 6



- Even though subjects with missing values shown in Figure 6 had no change or an increase, while all completers decrease, there is a systematic trend for slope to decrease as the response at Visit 1 decreases. By incorporating this trend into the model, we can avoid bias in estimates due to missing data. The missingness here is non-informative since everything we need in order to estimate average change can be obtained via the observed data.

- There are (at least) two modeling approaches that can be used for data in Figure 6 that assume a linear relationship between the slope and the starting value.

- One is to allow dependency of the outcome on starting value by including the baseline value as a covariate; the outcome can be modeled as change (Y_2–Y_1) or just the Y_2 value. In either case, the model is not 'longitudinal' any more, although we can still using

LMM methods in order to impute missing values and obtain fixed-effect estimates that adjust for the non-completers.

- The second modeling approach is to keep both measures as outcomes, add random intercept and slope for time, plus a covariance between the two. In particular, for the pattern in Fig. 6, there will be a negative covariance between the intercept and slope that allows for the dependency of the slope on the starting value.

- Unfortunately with only 2 time points, including a random slope for *visit* will likely lead to a mixed model fit with a non-positive-definite Hessian matrix and consequently some limited and/or questionable output. Below are fixed-effect estimates for the two approaches with the given data (using the change outcome for the first approach).

- Note that the estimates of the slope over time for both approaches are identical. However, to get the slope that includes non-completers for the first approach, a custom test needs to be constructed that incorporates the averages of all 7 subjects. The predicted values for subjects in Approach 2 are a bit wonky (perhaps related to the limited time points and NPD Hessian matrix), although the predicted values for the 3 non-completers are pretty accurate.

Modeling approach 1 (non-longitudinal):

```
data test3; input id visit y y_bl @@;
datalines;
1 1 11.5 11.5 1 2 9.5 11.5 2 1 10 10 2
2 8.7 10 3 1 8 8 3 2 7 8 4 1 7 7 4 2
6.5 7 5 1 4 4
5 2 . 4 6 1 3 3 6 2 . 3 7 1 2 2 7 2 . 2
;
data test3; set test3; v=visit;
 y_diff=y-y_bl;
```

Modeling approach 2 (longitudinal):

```
proc mixed data=test3; class id visit;
 model y= visit / solution outp=preddy;
 random intercept v / type=un
subject=id solution;
 lsmeans visit;
 estimate 'change' visit -1 1; run;
```



```
proc mixed data=test3; where visit=2;
 class id visit;
 model y_diff = y_bl / solution
outp=preddy;
 estimate 'slope, ave bl' intercept 1
y_bl 6.5; run;

Estimates

Label          Est    SE    DF t Value Pr>|t|
slope, ave bl -0.4031 0.1521 2 -2.65  0.1177
```

```
Convergence criteria met but final
hessian is not positive definite.

Estimates

Label    Est    SE   DF t Value  Pr>|t|
change -0.4031 0.4207 0  -0.96     .

Least Squares Means
Effect visit Est   SE  DF tValue Pr>|t|
visit    1 6.5000 1.3671 0 4.75     .
visit    2 6.0969 0.9546 0 6.39     .
```

- The modeling approaches mentioned above become more useful when there are t>2 time points; the 2nd (longitudinal) approach will likely not have fit issues once there are more time points to better estimate subject slopes over time, and tests will ‘work’, more likely having nonzero SE and DF values. The ‘baseline as covariate’ approach will also be more reasonable and the associated model is then truly longitudinal since there are at least t-1 (at least 2) time points to model as the outcome.

3.3 Missing data mechanisms

- In this subsection we discuss classes of missing data mechanisms that are commonly discussed in the literature, and with illustrations using the informative and non-informative dropout examples presented in the last subsection. Definitions for missing data mechanisms for given data are well defined in the literature, and include *missing completely at random* (MCAR), *missing at random* (MAR) and *missing not at random* (MNAR). Generally, definitions involve missing response (*Y*) data, while predictors (*X*) are considered complete. Such is the case for the formal definitions given in the text.

- For the simplest type of MCAR data, the probability that the response is missing is unrelated to any of the data, including the missing responses. Formally, this can be written as $P(M_{ij}=1|\mathbf{Y}_{i,obs},\mathbf{Y}_{i,miss},\mathbf{X}_i)=P(M_{ij}=1)$, where $\mathbf{Y}_{i,obs}$ and $\mathbf{Y}_{i,miss}$ denote the

observed and missing components of the responses, $\mathbf{X}_i$ denotes relevant predictors in the model, and M_{ij} is an indicator for missingness (1=missing, 0=observed), for subject *i* at time *j*.

- A slightly more general assumption is $P(M_{ij}=1|\mathbf{Y}_{i,obs},\mathbf{Y}_{i,miss},\mathbf{X}_i)=P(M_{ij}=1|\mathbf{X}_i)$, which is often called *covariate-dependent missingness*, but still generally considered a type of MCAR (see Hedeker and Gibbons, 2006; Fitzmaurice et al., 2011).

- Longitudinal data that do satisfy MCAR are much more likely to have covariate-dependent missingness rather than simplest type, such as when subjects tend to dropout more as time goes on, and hence requiring conditioning on *X*. But generally, MCAR is the most restrictive assumption and is probably the least likely to hold for real data.

- However, one can test whether data are MCAR or not fairly easily, while distinctions between other types of data missing mechanisms are more difficult if not impossible.

- The next level up is MAR data, which satisfies $P(M_{ij}=1|\mathbf{Y}_{i,obs},\mathbf{Y}_{i,miss},\mathbf{X}_i)=P(M_{ij}=1|\mathbf{Y}_{i,obs},\mathbf{X}_i)$. Modeling MAR data can still be done somewhat easily if the model contains the necessary variables for the observed data.

- As an example, consider progression of lung function over time for smokers and nonsmokers, where smokers have lower starting values and steeper drops in response over time compared with nonsmokers. Additionally smokers are more likely to dropout (40%, compared with 20% for nonsmokers). In this case the missingness depends on smoking status, but by including the relevant predictors in the model (*smoking status*, *time* and their interaction), we can accurately model the data.

- Note that within smoking status groups, the probability of missingness at Visit 2 is constant across subjects, and so whether or not a subject's response at Visit 2 is observed does not depend on its potentially unobserved value.

- The non-informative missing data example in Subsection 2.2 is also an example of MAR data as if the 'true' model for the data is similar to the sample of 7 subjects, where the expected Visit 2 response is a linear function of the Visit 1 response. In this case, once we have the Visit 1 response, we can obtain unbiased predicted values for Visit 2 regardless of whether the Visit 2 data were observed or unobserved.

- When the values of the missing data are related to the chance that they are missing (specifically, when the probability equation in the last paragraph does not hold), the mechanism is referred to as

missing not at random (MNAR; or in some places, termed ‘not missing at random’).

- For a simple example, consider a study where a health outcome is measured over time, where subjects are more likely to dropout once they become sick. If the health outcome measure decline for these sick subjects but we did not observe their outcomes during this state, then data are likely MNAR.

- The informative missing data example in Subsection 2.2 (re: Figure 5) is also most likely MNAR data. There is nothing that informs the analyst about what the values of missing data will be from observed data, and without more information, they would be unable to estimate either predicted trajectories for the 3 lower subjects or the average progression. The probability of missingness appears to be related not only to starting value, but progression, which depends on both Y_1 and Y_2.

- Unfortunately, there are no easy tests to determine whether data are MAR versus MNAR unless some additional information becomes available (e.g., some of the missing responses are randomly obtained). There are methods of estimation that do account for MNAR type of data, if there is concern that data may follow that, including pattern mixture models and selection models (e.g., see Diggle et al., 2002), and Kenward (1998) even suggested a selection model for 2-visit data with missing values. If there is enough uncertainty about MAR versus MNAR data, methods for the two approaches can always both be run in a ‘sensitivity fashion’ to help determine how much difference it makes.

- The example shown in Figures 3 and 4 is likely to be MAR data, since missing values appear to depend on observed responses, but not unobserved.

3.4 Inverse probability weighting (IPW)

See the course notes.

3.5 Simulations

See the course notes.

3.6 Computing EBLUPs for missing observations

See the course notes.

3.7 Approaches for missing X data

- When *Y* is missing at random (MAR) but covariate data are complete, then it is sufficient to use the standard linear mixed model in order to obtain unbiased estimates, as described above. However, when *X* is missing (potentially with some missing *Y*), standard likelihood based methods may not be sufficient.

- To address potential bias for missing *X* data, one might consider other likelihood-based algorithms, such as the EM algorithm, or another modeling approach, such as multiple imputation. Such approaches may be able to incorporate records that involve missing *X* data rather than just removing records. Although most standard procedures simply drop the records from analysis when covariate data are missing, there are ways to account for correlation between responses in such cases, as described above.

4 GEE and estimation in light of missing data

- When using GEE associated with GzLMs, unlike mixed models that employ more standard likelihood-based estimation methods, MAR-type data cannot necessary be handled by simply including key predictor variables. For GEE, a stronger assumption of MCAR is necessary in order to use typical estimation methods. Fitzmaurice, et al. (2011), also discuss how to employ weighting techniques for GEE models when data are MAR.

5 Preparation of data and specification of models in light of missing data

- In this section we focus on computational issues for data with missing values when fitting linear mixed models, and how you should specify a data set to get an accurate model fit when using computer software such as SAS or R.

- Note that which software you use makes a difference on the approach. Here we focus on data with serial correlation that can be modeled with an AR(1) or related structure. Including random effects such as a random intercept are less problematic in light of missing data since each pair of responses have the same model correlation, regardless of time between responses. On the other hand, the AR(1) is sensitive to the time between measurements, and so missing values need to be carefully considered.

5.1 Linear mixed models in SAS

- It is important to account for missing data in an analysis. For example, if we are fitting an AR(1) structure for the subject that is missing one response (out of 5), we need to know which value is missing so that we can correctly model the correlation. If the missing response is not either the first or last time point, then there will be a gap between the two time points that sandwich the

missing time point. We need to account for this gap by using ϕ^2 as the correlation between these times, rather than just ϕ.

- In order to account for missing responses (Y), then it is helpful to use an index variable for the repeated measures, such as

```
repeated time / subject=ID type=ar(1);
```

- The variable 'time' here must be a class variable. If you do not want to model time as a class variable, then you can simply create a new variable that is like the original time variable in every way except it is a class variable (e.g., call it *t*). Thus, you put *time* in the model but not the class statement, and you put *t* in the class statement and repeated statement but not the model statement. With this approach, you in fact do not need to include records with missing *Y* as long as all subjects are not missing *Y* for any one time point.

- If there is at least one time point for which all subjects have missing *Y*, all records should be included for all time points, and '.' entered as necessary when *Y* is missing.

- If you use this 'all records' approach, then you actually do not need to include the time index variable in the REPEATED statement and this will work fine as long as data are sorted by subject and then by time.

- An alternative here is to use a spatial structure, such as the spatial power structure (a.k.a. 'continuous AR(1)' structure). The variable used in the spatial context is defined as part of the structure type, so using an index variable becomes unnecessary.

- In the computation of estimates, the observed data and missing data are partitioned as previously described. Thus, $\mathbf{R}_{i,obs}$ for subject i would be an $r_i \times r_i$ matrix, where r_i represents the number of observed values for that subject.

- In the previous section, we only considered missing Y, not missing X. In SAS, PROC MIXED, if we have multiple predictor (X) variables, then a record is simply dropped from all analyses if it has at least one X variable with a missing value; these records are not used for estimation of parameters and no predicted values are computed for the associated Y, either.

- In this case, if we use the discrete AR(1) structure, it is important to include the index variable for repeated measures (e.g., time) in the REPEATED statement; just including a '.' for a missing X variable will not allow proper spacing for the correlation.

- Thus, it is important to do both of the following: (i) index the repeated measures by including 'time' in the REPEATED statement and (ii) use the 'all records' approach, placing '.' when a variable is missing, whether it is X or Y or both.

- However, if X (for a particular predictor) is missing for all subjects at a given time point, then that time point is not factored into the analysis and the covariances will not be modeled properly for structures such as AR(1). In this case, I would suggest using the continuous AR(1) or other spatial structure, if possible. If there are relatively few times points, the UN structure is another possibility.

- In PROC MIXED, if we use a spatial covariance structure instead of the AR(1) structure, we do not need to keep records where the response is missing (both between observation periods as well as within). This is because we define a time variable in the spatial structure (e.g., 'date') to indicate how far the observations are apart, which then determines the strength of the covariance between measurements.

5.2 Linear mixed models in R

- When specifying serial correlation in a model with no random effects, the *gls* function can be used in R. Recall that missing values are specified using 'NA'. Some functions such as *gls* or *lme* (from the *nlme* package) cannot process the records with 'NA' without more instruction about how to deal with them. Specifically, telling the function to omit or exclude the records will allow the model to fit.

- Unfortunately, for the discrete AR(1) structure, information about spacing will be lost. Thus, using the continuous AR(1) structure is suggested here. Since records are dropped, we can retain information about correct spacing by including the variable that specifies when observations were taken.

- To illustrate this, we consider the first 5 male subjects from the Orthodont data (from R library). We purposely create missing values for 2 responses as shown below. A comparison with SAS is given to the right. Note that in SAS we can use either the discrete or continuous AR(1). The discrete AR(1) in R is specified by 'corAR(1)' while the continuous AR(1) is specified by 'corCAR(1)'. The latter is the same as the spatial power structure in SAS.

- As before, there are some differences in what is presented in R and SAS output. Also, the correlation parameter estimate differs a bit depending on whether the discrete or continuous AR(1) approach is used. Specifically, for the discrete AR(1), age is treated categorically, so that the reported correlation is relevant for two adjacent levels (e.g., 8 and 10 years). When the continuous AR(1) is used, the correlation is relevant for one unit in the time-indexing variable (i.e., age). Thus, to get the correlation between a 2-year gap in ages, the estimate is squared (ϕ^{days}: $0.8638^2 = 0.7462$).

```
Data:

obs   distance   age   Subject   Sex
1       26.0       8     M01     Male
2       25.0      10     M01     Male
3       NA        12     M01     Male
4       31.0      14     M01     Male
5       21.5       8     M02     Male
6       22.5      10     M02     Male
7       23.0      12     M02     Male
8       26.5      14     M02     Male
```

```
Data:  same as to left, although missing
values are specified by '.' Rather than 'NA'.

Approach 1:  discrete AR(1)

proc mixed data=ortho method=reml;
class subject;
model distance = age / solution;
repeated / type=AR(1) subject=subject; run;

Covariance Parameter Estimates
```



```
9       23.0       8     M03     Male
10      22.5      10     M03     Male
11      24.0      12     M03     Male
12      27.5      14     M03     Male
13      25.5       8     M04     Male
14      27.5      10     M04     Male
15      26.5      12     M04     Male
16      27.0      14     M04     Male
17      20.0       8     M05     Male
18      NA        10     M05     Male
19      22.5      12     M05     Male
20      26.0      14     M05     Male

Continuous AR(1) approach:

fm3=gls(y~age,
correlation=corCAR1(form=~age|Subject),
na.action=na.omit)

> fm3
Generalized least squares fit by REML
  Model: y ~ age
  Data: NULL
  Log-restricted-likelihood: -35.38116

Coefficients:
(Intercept)         age
 17.1304025   0.7311602
```

```
Cov Parm   Subject   Estimate
AR(1)      Subject    0.7462
Residual              5.7697

Fit Statistics
-2 Res Log Likelihood   70.8
AIC (smaller is better) 74.8

Solution for Fixed Effects
Effect    Estimate  SE     DF t Value  Pr>|t|
Intercept 17.1304  2.3105   4   7.41   0.0018
age        0.7312  0.1936  12   3.78   0.0026

Approach 2:  continuous AR(1) (i.e., spatial
power)

proc mixed data=ortho method=reml;
class subject;
model distance = age / solution;
repeated / type=sp(pow)(age) subject=subject;
run;

Solution for Fixed Effects
Effect    Estimate  SE    DF  t Value  Pr>|t|
Intercept 17.1304 2.3105   4   7.41    0.0018
age        0.7312 0.1936  12   3.78    0.0026
```

Correlation relevant for responses 2 years apart.

```
Correlation Structure: Continuous AR(1)
 Formula: ~age | Subject
 Parameter estimate(s):
      Phi
0.8638075
Degrees of freedom: 18 total; 16 residual
Residual standard error: 2.402016
```

Correlation relevant for responses 1 year apart.

```
Covariance Parameter Estimates
Cov Parm   Subject   Estimate
SP(POW)    Subject    0.8638
Residual              5.7698

Fit Statistics
-2 Res Log Likelihood   70.8
AIC (smaller is better) 74.8
```

- In R, the same issues apply when X is missing. Again, the easiest approach is to just use the continuous AR(1) structure.

5.3 GEEs with SAS and R

- In SAS, PROC GENMOD, the AR(1) working structure is available to model repeated measures over time. If responses are unequally spaced, unfortunately spatial structures are not available to use. However, for most data it is possible to get around this issue by creating equal time units and then filling in the data set with missing values, as necessary. This will work even when covariate measures, other than time, are missing too.

- For example, say data are collected on weekdays but the response and covariates other than time are not collected on weekends. In the data set, just include a record for every day in the month, and put missing values in for *Y* and covariates on the weekend days (other than for ‘day’, which should be complete), and employ the AR(1) working structure for GEE. The fitted model will reflect the unequal spacing caused by no collection on weekends. R has at least a couple of packages to fit GEE models: the *geeglm* function within the *geepack* package is one route, and the *gee* function within the *gee* package is another. However, within the default settings the correct spacing cannot be specified when there are missing data for either approach.

5.4 Summary – modeling serial correlation in light of missing data

- The AR(1) covariance structure is ideal for many data sets where serial correlation is involved. The standard AR(1) covariance structure is defined for 'discrete time' data and is most appropriate for data collected at equal intervals.

- If data are not collected at equal intervals, if subjects vary in times of measurement, or for data with missing values, a spatial covariance structure (e.g., spatial power or 'continuous AR(1)) can be used in mixed models to get accurate results. In fact, it should work fine even for discrete data; it is really just a more general structure, with standard AR(1) as a special case. The only exception would be if for some reason convergence cannot be obtained via the spatial structure, but can be with the AR(1) structure.

- For GEE models, spatial structures are not available, so for serial correlation usually the standard AR(1) structure is the best bet, but data must include records with missing values.

6 Case study: data with more than two responses

- The examples so far in this section have involved 2 repeated measures over time. Most longitudinal data will have more time points. Although this will probably complicate the analysis, it should allow the researcher to better understand whether missing data is an issue, and also to account for it in the modeling approach.

- Application: recall the eNO data first presented in the Graphs chapter. Below are the data, which now includes the 6mo and 1yr time points (eno_1=pre challenge; eno_2=post challenge; eno_3=6 months after challenge; eno_4=1 year after challenge).

- Note that there is technically only 1 day difference between the pre and post challenge measurements. However, to allow for visual interpretation, a spread of 10 days was used between eno_1 and eno_2; otherwise data were plotted metrically for time. In the data, missing values were represented by '.' (For R, they would be represented with 'NA'.)

- The data were sorted by eno_1 and demonstrate that those with higher baseline eNO (indicating more inflammation) were more likely to drop out later on, although the subject with the highest starting eNO and biggest reaction to aspirin was a completer.

- A straight mean of available data shows an increase in eNO after the aspirin challenge, which then drops somewhat at 6 months and 1 year. This is also apparent in the graph of individual subjects. Although dropouts tend to occur as time goes on, there are a few cases where subjects missed intermediate time points but actually came back. These are represented with dashed lines (they both missed the 6mo time point).

- Two questions for the reader: (1) Based on what you see, what type of missing data mechanism would you expect the data to follow? (2) How would you check for and handle (if applicable) the missing data?

ID	eno_1	eno_2	eno_3	eno_4
1	9.6	17.7	.	35.1
2	11.1	11.7	35.4	.
3	16.4	11.6	30.3	17.0
4	20.8	19.3	11.4	.
5	22.3	20.7	16.3	24.7
6	24.4	37.9	.	.
7	25.8	17.4	17.1	21.1
8	27.0	17.8	34.9	30.5
9	27.7	41.2	43.0	.
10	31.8	24.3	24.4	15.4
11	34.1	55.4	53.5	.
12	36.9	60.7	32.4	.
13	37.4	42.5	62.7	67.0
14	38.2	.	.	.
15	47.4	79.8	97.1	.
16	51.7	90.4	.	.
17	53.8	67.4	33.3	.
18	57.9	85.7	.	.
19	69.6	116.8	.	.
20	80.5	79.1	43.3	.
21	114.7	176.3	.	103.6