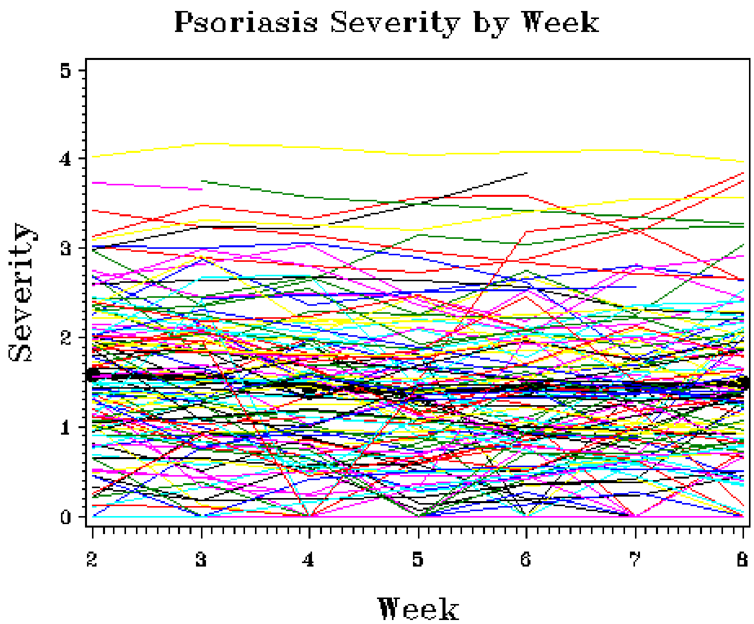
Example 4: Within-Person Fluctuation in Symptom Severity over Time
(complete syntax and output available for SAS, STATA, and R electronically)



These are real data from a study of weekly fluctuation in psoriasis severity. There was no intervention and no real reason to expect systematic growth (as shown by the mostly flat solid black line for the means).

But we still need ensure good fit for the model for the variance. The variances across occasions may need to differ, and the covariances across occasions may need to differ as well.

We will include just a fixed intercept in the model for the means and examine different ACS models for the variance, including R-only and G+R models for the variance.

To begin, let's see what the observed pattern of variances and covariances over time looks like by estimating an *n*-order unstructured **R** baseline model for the variance → each variance and covariance estimated is separately, with no constraints, so this is a description, not a predictive model. It is equivalent to the "H1" model in SEM terminology with respect to the occasions' variances and covariances, but not with respect to the occasion means (we have only estimated one mean here instead of all possible means, as in the H1 model). Note: you can only estimate an unstructured model if time is balanced across persons and you have more people than parameters estimated!

n-order Unstructured (UN) R-only Model: the answer key of all possible variances and covariances

```
TITLE1 "SAS n-order Unstructured (UN) R-only Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity =  / SOLUTION DDFM=Satterthwaite;
     REPEATED week / TYPE=UN R RCORR SUBJECT=subid;
     ODS OUTPUT InfoCrit=FitUN; * Save for LRT;
RUN; TITLE1;

display "STATA n-order Unstructured R-only Model"
mixed severity ,  || subid:  , noconstant           ///
         variance reml residuals(unstructured,t(week)) ///
         dfmethod(satterthwaite) dftable(pvalue)
      estat wcorrelation, covariance  // R matrix
      estat wcorrelation              // RCORR matrix
      estat ic, n(124)                // AIC and BIC
      estimates store FitUN           // Save for LRT

# It insists I make a new week variable starting at 1, not 2 (insert eye-roll)
Example4$week1=Example4$week-1

print("R GLS n-order Unstructured R-only Model")
UN = gls(data=Example4, method="REML", model=severity~1,
         correlation=corSymm(form=~as.numeric(week1)|subid),
         weights=varIdent(form=~1|week1))
print("Show results using incorrect DDF with total leftover variance")
print("Total variance per occasion is created using SD multiplier")
summary(UN); summary(UN)$sigma^2
print("Show R and RCORR matrices for first person")
getVarCov(UN, individual="100", type="marginal")
corMatrix(UN$modelStruct$corStruct)[[6]]
```

SAS Output:

Note that given n=7, this model requires $n*(n+1) / 2$ = 28 covariance parameters!! You'd need at least 28 people to estimate it (and it may not be possible even then).

Dimensions

Covariance Parameters	28	Cov Parms = total number of parameters in model for variances
Columns in X	1	Columns in X = total number of fixed effects plus intercept
Columns in Z	0	Columns in Z = total number of U's (not counting covariances)
Subjects	124	Subjects = number of persons in level 2 (with at least 1 obs)
Max Obs Per Subject	7	Max Obs Per Subject = max number of time points per person

Estimated R Matrix for SUBID 100 → TOTAL (marginal) variance and covariance → ANSWER KEY

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.7477**	0.7190	0.6817	0.6807	0.6124	0.6182	0.6328
2	0.7190	**0.8266**	0.7428	0.7475	0.6571	0.6619	0.6541
3	0.6817	0.7428	**0.8346**	0.7498	0.6524	0.6390	0.6172
4	0.6807	0.7475	0.7498	**0.8366**	0.7292	0.7206	0.6805
5	0.6124	0.6571	0.6524	0.7292	**0.8285**	0.7280	0.7313
6	0.6182	0.6619	0.6390	0.7206	0.7280	**0.7911**	0.7352
7	0.6328	0.6541	0.6172	0.6805	0.7313	0.7352	**0.8464**

Estimated R Correlation Matrix for SUBID 100 → TOTAL (marginal) correlation → answer key

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.9146	0.8630	0.8608	0.7781	0.8038	0.7955
2	**0.9146**	1.0000	0.8943	0.8989	0.7940	0.8185	0.7820
3	**0.8630**	**0.8943**	1.0000	0.8974	0.7846	0.7864	0.7343
4	**0.8608**	**0.8989**	**0.8974**	1.0000	0.8760	0.8859	0.8088
5	**0.7781**	**0.7940**	**0.7846**	**0.8760**	1.0000	0.8993	0.8734
6	**0.8038**	**0.8185**	**0.7864**	**0.8859**	**0.8993**	1.0000	0.8985
7	**0.7955**	**0.7820**	**0.7343**	**0.8088**	**0.8734**	**0.8985**	1.0000

Covariance Parameter Estimates table is too big to print—it has 28 entries!

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
893.1	**28**	**949.1**	951.3	981.2	**1028.1**	1056.1

The fit statistics for this n-order unstructured **R**-only model will serve as a baseline with which to compare more parsimonious models for the variance.

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.5178	0.07429	122	20.43	<.0001

Ok, so this is what we are trying to reproduce… next we will try a few **R**-only models and see how they fit relative to this Unstructured answer key—can we use fewer parameters to capture these patterns and not decrease model fit?

<u>Compound Symmetry (CS) R-only Model</u>: equal total variances and covariances across weeks

```
TITLE1 "SAS Compound Symmetry (CS) R-only Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     REPEATED week / TYPE=CS R RCORR SUBJECT=subid;
     ODS OUTPUT InfoCrit=FitCS; * Save for LRT;
RUN; TITLE1;
* Model comparison: not worse than UN?;
  %FitTest(FitFewer=FitCS, FitMore=FitUN);
```

```
display "STATA Compound Symmetry (CS) R-only Model"
mixed severity ,  || subid:  , noconstant              ///
         variance reml residuals(exchangeable,t(week)) ///
         dfmethod(satterthwaite) dftable(pvalue)
      estat wcorrelation, covariance  // R matrix
      estat wcorrelation              // RCORR matrix
      estat ic, n(124)                // AIC and BIC
      estimates store FitCS           // Save for LRT
      lrtest FitUN FitCS              // Worse than UN?
```

```
print("R GLS Compound Symmetry R-only Model")
CS = gls(data=Example4, method="REML", model=severity~1,
         correlation=corCompSymm(form=~1|subid))
print("Show results using incorrect DDF"); summary(CS)
print("Show R and RCORR matrices for first person")
getVarCov(CS, individual="100", type="marginal");
corMatrix(CS$modelStruct$corStruct)[[6]]
print("LRT: Is CS worse than UN?"); anova(UN,CS)
```

SAS Output:

Estimated R Matrix for SUBID 100 → TOTAL (marginal) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	0.8127	0.6820	0.6820	0.6820	0.6820	0.6820	0.6820
2	0.6820	0.8127	0.6820	0.6820	0.6820	0.6820	0.6820
3	0.6820	0.6820	0.8127	0.6820	0.6820	0.6820	0.6820
4	0.6820	0.6820	0.6820	0.8127	0.6820	0.6820	0.6820
5	0.6820	0.6820	0.6820	0.6820	0.8127	0.6820	0.6820
6	0.6820	0.6820	0.6820	0.6820	0.6820	0.8127	0.6820
7	0.6820	0.6820	0.6820	0.6820	0.6820	0.6820	0.8127

Estimated R Correlation Matrix for SUBID 100 → TOTAL (marginal) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8393	0.8393	0.8393	0.8393	0.8393	0.8393
2	0.8393	1.0000	0.8393	0.8393	0.8393	0.8393	0.8393
3	0.8393	0.8393	1.0000	0.8393	0.8393	0.8393	0.8393
4	0.8393	0.8393	0.8393	1.0000	0.8393	0.8393	0.8393
5	0.8393	0.8393	0.8393	0.8393	1.0000	0.8393	0.8393
6	0.8393	0.8393	0.8393	0.8393	0.8393	1.0000	0.8393
7	0.8393	0.8393	0.8393	0.8393	0.8393	0.8393	1.0000

The intraclass correlation in RCORR can be computed as:

$$\text{ICC} = \frac{0.6820}{0.6820 + 0.1306} = .8393$$

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
CS	SUBID	0.6820	0.09019	7.56	<.0001 Compound symmetry parameter
Residual		0.1306	0.007270	17.97	<.0001 Total variance

```
                         Information Criteria
Neg2LogLike     Parms        AIC         AICC        HQIC         BIC        CAIC
     1049.7         2     1053.7       1053.7      1056.0      1059.4      1061.4

  Null Model Likelihood Ratio Test
    DF     Chi-Square      Pr > ChiSq
     1         950.14          <.0001
```

This default output is only useful in this case: Is the CS model better than an e-only model? Yes, $-2\Delta LL(1) = 950.2$, $p < 001$

```
                       Solution for Fixed Effects
                             Standard
Effect          Estimate        Error      DF    t Value     Pr > |t|
Intercept         1.4767      0.07548     123      19.56       <.0001

Likelihood Ratio Test for FitCS vs. FitUN
         Neg2Log
Name      Like      Parms        AIC         BIC      DevDiff      DFdiff      Pvalue
FitCS     1049.7       2      1053.7      1059.4         .            .           .
FitUN      893.1      28       949.1      1028.1      156.585        26           0
```

Is CS worse than UN? Yes! $-2\Delta LL(26) = 156.6$, $p < 001$

<u>Auto-Regressive (AR) R-only Model</u>: equal total variances and an AR1 total correlation across weeks

```
TITLE1 "SAS Auto-Regressive (AR) R-only Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     REPEATED week / TYPE=AR(1) R RCORR SUBJECT=subid;
     ODS OUTPUT InfoCrit=FitAR; * Save for LRT;
RUN; TITLE1;
* Model comparison: not worse than UN?;
  %FitTest(FitFewer=FitAR, FitMore=FitUN);
```

For unbalanced time, you can use `TYPE=SP(POW)(week)` instead

```
display "STATA Auto-Regressive (AR) R-only Model"
mixed severity ,  || subid:  , noconstant         ///
         variance reml residuals(ar1,t(week))     ///
         dfmethod(satterthwaite) dftable(pvalue)
      estat wcorrelation, covariance      // R matrix
      estat wcorrelation                  // RCORR matrix
      estat ic, n(124)                    // AIC and BIC
      estimates store FitAR               // Save for LRT
      lrtest FitUN FitAR                  // Worse than UN?
```

Option for unbalanced time are not available in STATA as far as I know

```
print("R GLS Auto-Regressive R-only Model")
AR = gls(data=Example4, method="REML", model=severity~1,
         correlation=(corAR1(form=~as.numeric(week1)|subid)))
print("Show results using incorrect DDF"); summary(AR)
print("Show R and RCORR matrices for first person")
getVarCov(AR, individual="100", type="marginal");
corMatrix(AR$modelStruct$corStruct)[[6]]
print("LRT: Is AR worse than UN?"); anova(UN,AR)
```

For unbalanced time, you can use `corCAR1` instead

SAS Output:

Estimated R Matrix for SUBID 100 → TOTAL (marginal) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.7999**	0.7163	0.6415	0.5745	0.5145	0.4607	0.4126
2	0.7163	**0.7999**	0.7163	0.6415	0.5745	0.5145	0.4607
3	0.6415	0.7163	**0.7999**	0.7163	0.6415	0.5745	0.5145
4	0.5745	0.6415	0.7163	**0.7999**	0.7163	0.6415	0.5745
5	0.5145	0.5745	0.6415	0.7163	**0.7999**	0.7163	0.6415
6	0.4607	0.5145	0.5745	0.6415	0.7163	**0.7999**	0.7163
7	0.4126	0.4607	0.5145	0.5745	0.6415	0.7163	**0.7999**

```
            Estimated R Correlation Matrix for SUBID 100 → TOTAL (marginal) correlation
 Row        Col1        Col2        Col3        Col4        Col5        Col6        Col7
   1      1.0000      0.8955      0.8020      0.7182      0.6432      0.5760      0.5158
   2      0.8955      1.0000      0.8955      0.8020      0.7182      0.6432      0.5760
   3      0.8020      0.8955      1.0000      0.8955      0.8020      0.7182      0.6432
   4      0.7182      0.8020      0.8955      1.0000      0.8955      0.8020      0.7182
   5      0.6432      0.7182      0.8020      0.8955      1.0000      0.8955      0.8020
   6      0.5760      0.6432      0.7182      0.8020      0.8955      1.0000      0.8955
   7      0.5158      0.5760      0.6432      0.7182      0.8020      0.8955      1.0000

                    Covariance Parameter Estimates
                                         Standard         Z
Cov Parm     Subject     Estimate       Error       Value        Pr Z
AR(1)        SUBID         0.8955     0.01148       78.01      <.0001 Total auto-regressive correlation
Residual                   0.7999     0.08044        9.94      <.0001 Total variance (equal across weeks)

                              Information Criteria
Neg2LogLike      Parms        AIC        AICC        HQIC         BIC        CAIC
      990.9          2      994.9       994.9       997.2      1000.5      1002.5

                       Solution for Fixed Effects
                             Standard
Effect         Estimate         Error      DF     t Value     Pr > |t|
Intercept        1.5148       0.07031     129       21.54       <.0001
```

AR1 also forces all variances equal, but with correlations declining sharply with time (r, r^2, r^3... = .8955, .8020, .7182…). Relative to the UN model, the BIC (but not the AIC) is happier.

```
Likelihood Ratio Test for FitAR1 vs. FitUN
          Neg2Log
 Name       Like       Parms        AIC         BIC     DevDiff     DFdiff        Pvalue
FitAR1     990.9          2       994.9      1000.5        .           .               .
FitUN      893.1         28       949.1      1028.1     97.7472        26      3.0326E-10
```

Is AR1 worse than UN? Yes! $-2\Delta LL(26) = 97.7$, $p < 001$

Toeplitz(*n*) R-only Model: equal total variances and 6 lagged total covariances across weeks

```
TITLE1 "SAS Toeplitz (n=7 bands) R-only Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     REPEATED week / TYPE=TOEP(7) R RCORR SUBJECT=subid;
     ODS OUTPUT InfoCrit=FitToep7; * Save for LRT;
RUN; TITLE1;
* Model comparison: not worse than UN?;
  %FitTest(FitFewer=FitToep7, FitMore=FitUN);
```

For Toeplitz in SAS, (*n*) indicates the total number of bands including the residual variance on the diagonal + 6 covariances for the 7 occasions.

For Toeplitz in STATA, (*n*) indicates the number of bands NOT including the residual variance on the diagonal: just 6 covariances for the 7 occasions.

```
display "STATA Toeplitz (n=7 bands) R-only Model"
mixed severity ,  || subid:  , noconstant              ///
         variance reml residuals(toeplitz6,t(week)) ///
         dfmethod(satterthwaite) dftable(pvalue)
      estat wcorrelation, covariance    // R matrix
      estat wcorrelation                // RCORR matrix
      estat ic, n(124)                  // AIC and BIC
      estimates store FitToep7          // Save for LRT
      lrtest FitUN FitToep7             // Worse than UN?
```

```
# Toeplitz is not a pre-defined structure in R gls
```

SAS Output:

Estimated R Matrix for SUBID 100 → TOTAL (marginal) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.8103**	0.7257	0.6974	0.6576	0.6394	0.6368	0.6541
2	0.7257	**0.8103**	0.7257	0.6974	0.6576	0.6394	0.6368
3	0.6974	0.7257	**0.8103**	0.7257	0.6974	0.6576	0.6394
4	0.6576	0.6974	0.7257	**0.8103**	0.7257	0.6974	0.6576
5	0.6394	0.6576	0.6974	0.7257	**0.8103**	0.7257	0.6974
6	0.6368	0.6394	0.6576	0.6974	0.7257	**0.8103**	0.7257
7	0.6541	0.6368	0.6394	0.6576	0.6974	0.7257	**0.8103**

Estimated R Correlation Matrix for SUBID 100 → TOTAL (marginal) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8955	0.8606	0.8115	0.7890	0.7859	0.8072
2	**0.8955**	1.0000	0.8955	0.8606	0.8115	0.7890	0.7859
3	**0.8606**	**0.8955**	1.0000	0.8955	0.8606	0.8115	0.7890
4	**0.8115**	**0.8606**	**0.8955**	1.0000	0.8955	0.8606	0.8115
5	**0.7890**	**0.8115**	**0.8606**	**0.8955**	1.0000	0.8955	0.8606
6	**0.7859**	**0.7890**	**0.8115**	**0.8606**	**0.8955**	1.0000	0.8955
7	**0.8072**	**0.7859**	**0.7890**	**0.8115**	**0.8606**	**0.8955**	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
TOEP(2)	SUBID	0.7257	0.09004	8.06	<.0001	Total Covariance for t-1
TOEP(3)	SUBID	0.6974	0.08997	7.75	<.0001	Total Covariance for t-2
TOEP(4)	SUBID	0.6576	0.09003	7.30	<.0001	Total Covariance for t-3
TOEP(5)	SUBID	0.6394	0.09019	7.09	<.0001	Total Covariance for t-4
TOEP(6)	SUBID	0.6368	0.09066	7.02	<.0001	Total Covariance for t-5
TOEP(7)	SUBID	0.6541	0.09179	7.13	<.0001	Total Covariance for t-6
Residual		0.8103	0.09014	8.99	<.0001	Total variance (equal across weeks)

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
926.0	7	940.0	940.1	948.0	959.7	966.7

The TOEP(n) model fits almost not worse than the n-order UN model, $-2\Delta LL(21) = 32.9$, $p = .047$ (and AIC and BIC agree).

Could this be our model? And does it need to have heterogeneous total variances across weeks?

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.4884	0.07533	122	19.76	<.0001

Likelihood Ratio Test for FitToep7 vs. FitUN

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitToep7	926.0	7	940.0	959.7	.	.	.
FitUN	893.1	28	949.1	1028.1	32.8656	21	0.047729

<u>Heterogeneous Toeplitz(*n*) R-only Model</u>: Unequal total variances and 6 lagged total correlations

```
TITLE1 "SAS Heterogeneous Toeplitz (n=7 bands) R-only Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     REPEATED week / TYPE=TOEPH(7) R RCORR SUBJECT=subid;
     ODS OUTPUT InfoCrit=FitToep7H; * Save for LRT;
RUN; TITLE1;
* Model comparison: did het variance help?;
  %FitTest(FitFewer=FitToep7, FitMore=FitToep7H);
* Model comparison: now not worse than UN?;
  %FitTest(FitFewer=FitToep7H, FitMore=FitUN);
```

This model is only available directly in SAS MIXED as far as I can tell.

SAS Output:

Estimated R Matrix for SUBID 100 → TOTAL (marginal) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.7346**	0.6916	0.6896	0.6409	0.6363	0.5985	0.6332
2	0.6916	**0.8077**	0.7512	0.7126	0.6866	0.6322	0.6485
3	0.6896	0.7512	**0.8665**	0.7667	0.7540	0.6738	0.6766
4	0.6409	0.7126	0.7667	**0.8414**	0.7718	0.7039	0.6861
5	0.6363	0.6866	0.7540	0.7718	**0.8781**	0.7470	0.7431
6	0.5985	0.6322	0.6738	0.7039	0.7470	**0.7882**	0.7314
7	0.6332	0.6485	0.6766	0.6861	0.7431	0.7314	**0.8416**

Total *covariances are unequal* within a time lag band (because the total variances are allowed to differ over weeks), but total *correlations are equal* within a time lag band.

Estimated R Correlation Matrix for SUBID 100 → TOTAL (marginal) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8979	0.8644	0.8153	0.7923	0.7865	0.8053
2	**0.8979**	1.0000	0.8979	0.8644	0.8153	0.7923	0.7865
3	0.8644	**0.8979**	1.0000	0.8979	0.8644	0.8153	0.7923
4	0.8153	0.8644	**0.8979**	1.0000	0.8979	0.8644	0.8153
5	0.7923	0.8153	0.8644	**0.8979**	1.0000	0.8979	0.8644
6	0.7865	0.7923	0.8153	0.8644	**0.8979**	1.0000	0.8979
7	0.8053	0.7865	0.7923	0.8153	0.8644	**0.8979**	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
Var(1)	SUBID	0.7346	0.09397	7.82	<.0001	Total variance at week 1
Var(2)	SUBID	0.8077	0.1016	7.95	<.0001	Total variance at week 2
Var(3)	SUBID	0.8665	0.1129	7.68	<.0001	Total variance at week 3
Var(4)	SUBID	0.8414	0.1115	7.55	<.0001	Total variance at week 4
Var(5)	SUBID	0.8781	0.1162	7.56	<.0001	Total variance at week 5
Var(6)	SUBID	0.7882	0.1030	7.65	<.0001	Total variance at week 6
Var(7)	SUBID	0.8416	0.1092	7.70	<.0001	Total variance at week 7
TOEPH(1)	SUBID	0.8979	0.01292	69.51	<.0001	Total correlation for t-1
TOEPH(2)	SUBID	0.8644	0.01744	49.57	<.0001	Total correlation for t-2
TOEPH(3)	SUBID	0.8153	0.02429	33.56	<.0001	Total correlation for t-3
TOEPH(4)	SUBID	0.7923	0.02790	28.40	<.0001	Total correlation for t-4
TOEPH(5)	SUBID	0.7865	0.03056	25.73	<.0001	Total correlation for t-5
TOEPH(6)	SUBID	0.8053	0.03340	24.11	<.0001	Total correlation for t-6

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
921.5	13	947.5	947.9	962.3	984.1	997.1

Relative to the homogeneous Toeplitz: $-2\Delta LL(6) = 4.5$, $p = .604$. Separate variances don't improve model fit (AIC and BIC agree). The TOEPH(n) model also fits significantly worse than the UN(n) model, $-2\Delta LL(15) = 28.3$, $p = .020$.

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.5204	0.07417	124	20.50	<.0001

Likelihood Ratio Test for FitToep7 vs. FitToep7H

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitToep7	926.0	7	940.0	959.7	.	.	.
FitToep7H	921.5	13	947.5	984.1	4.54147	6	0.60381

Likelihood Ratio Test for FitToep7H vs. FitUN

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitToep7H	921.5	13	947.5	984.1	.	.	.
FitUN	893.1	28	949.1	1028.1	28.3241	15	0.019626

Another set of possible models: Just as we'll see later in models for within-person change, models for within-person fluctuation can make use of a combination of the **G** matrix (for between-person random effects) and the **R** matrix (for within-person residuals) to recreate the total (marginal) variance-covariance matrix. Adding a random intercept variance in **G** removes the primary source of variance and covariance from the **R** matrix, such that it will be easier to find a model for what is left in **R**. First, let's see the pattern of just the RESIDUAL WP variances and covariances…

Random Intercept in G + UN(*n*−1) R Model → Have to eliminate last covariance for the model to be identified because there is only one lag6 covariance (T1 with T7), so lag6 is not separately estimable

```
TITLE1 "SAS Random Intercept + Unstructured(n-1) R Model";
TITLE2 "Equal to n-Order Unstructured R-only Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     RANDOM INTERCEPT / TYPE=UN V VCORR G SUBJECT=subid;
     REPEATED week / TYPE=UN(6) R RCORR SUBJECT=subid;
RUN; TITLE1; TITLE2;

display "STATA Random Intercept + Unstructured(n-1) R Model"
display "Equal to n-Order Unstructured R-only Model"
mixed severity ,  || subid:  ,                         ///
         variance reml residuals(banded6,t(week)) ///
         dfmethod(satterthwaite) dftable(pvalue)
      estat recovariance, relevel(subid) // G matrix
      estat wcorrelation, covariance      // V matrix
      estat wcorrelation                  // VCORR matrix
      estat ic, n(124)                    // AIC and BIC

# Cannot see how to fit UN6 R matrix
# using R lme (defaults to UN7)
```

SAS now has a random statement that specifies the **G** matrix to have a random intercept variance.

The **V** matrix shows the combined (total, marginal) predicted matrix given by **R** (conditional) and **G**.

The residual correlations in RCORR are minimal after removing the contribution of the random intercept variance, such that **R** may be able to have a simpler structure yet still fit well.

SAS Output:

Estimated R Matrix for SUBID 100 → WP residual (conditional) variances and covariances

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1148**	0.08620	0.04892	0.04793	-0.02044	-0.01461	
2	0.08620	**0.1938**	0.1100	0.1147	0.02429	0.02907	0.02126
3	0.04892	0.1100	**0.2018**	0.1170	0.01962	0.006156	-0.01561
4	0.04793	0.1147	0.1170	**0.2038**	0.09643	0.08783	0.04774
5	-0.02044	0.02429	0.01962	0.09643	**0.1957**	0.09523	0.09853
6	-0.01461	0.02907	0.006156	0.08783	0.09523	**0.1583**	0.1024
7		0.02126	-0.01561	0.04774	0.09853	0.1024	**0.2136**

Estimated R Correlation Matrix for SUBID 100 → WP residual (conditional) correlations

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.5778	0.3213	0.3133	-0.1364	-0.1084	
2	0.5778	1.0000	0.5563	0.5773	0.1247	0.1660	0.1045
3	0.3213	0.5563	1.0000	0.5771	0.09875	0.03445	-0.07519
4	0.3133	0.5773	0.5771	1.0000	0.4830	0.4891	0.2288
5	-0.1364	0.1247	0.09875	0.4830	1.0000	0.5412	0.4820
6	-0.1084	0.1660	0.03445	0.4891	0.5412	1.0000	0.5571
7		0.1045	-0.07519	0.2288	0.4820	0.5571	1.0000

Estimated G Matrix → BP random intercept variance

Row	Effect	subject id	Col1
1	Intercept	100	0.6328

This random intercept variance exactly matches the last lag covariance (7,1) from the *n*-order UN **R**-only model.

```
Estimated V Matrix for SUBID 100 → TOTAL (marginal) variance and covariance (matches Unstructured R-only)
 Row       Col1      Col2      Col3      Col4      Col5      Col6      Col7
   1     0.7477    0.7190    0.6817    0.6807    0.6124    0.6182    0.6328
   2     0.7190    0.8266    0.7428    0.7475    0.6571    0.6619    0.6541
   3     0.6817    0.7428    0.8346    0.7498    0.6524    0.6390    0.6172
   4     0.6807    0.7475    0.7498    0.8366    0.7292    0.7206    0.6805
   5     0.6124    0.6571    0.6524    0.7292    0.8285    0.7280    0.7313
   6     0.6182    0.6619    0.6390    0.7206    0.7280    0.7911    0.7352
   7     0.6328    0.6541    0.6172    0.6805    0.7313    0.7352    0.8464

   Estimated V Correlation Matrix for SUBID 100 → TOTAL correlation (matches Unstructured R-only RCORR)
 Row       Col1      Col2      Col3      Col4      Col5      Col6      Col7
   1     1.0000    0.9146    0.8630    0.8608    0.7781    0.8038    0.7955
   2     0.9146    1.0000    0.8943    0.8989    0.7940    0.8185    0.7820
   3     0.8630    0.8943    1.0000    0.8974    0.7846    0.7864    0.7343
   4     0.8608    0.8989    0.8974    1.0000    0.8760    0.8859    0.8088
   5     0.7781    0.7940    0.7846    0.8760    1.0000    0.8993    0.8734
   6     0.8038    0.8185    0.7864    0.8859    0.8993    1.0000    0.8985
   7     0.7955    0.7820    0.7343    0.8088    0.8734    0.8985    1.0000

                             Information Criteria
Neg2LogLike      Parms        AIC       AICC       HQIC        BIC       CAIC
      893.1         28      949.1      951.3      981.2     1028.1     1056.1

                     Solution for Fixed Effects
                          Standard
Effect         Estimate      Error       DF    t Value    Pr > |t|
Intercept        1.5178    0.07429      122      20.43      <.0001
```

We can specify a model with constant residual variance (e-only) in the **R** matrix as our most simple baseline for what should be in **R**. Accordingly, our output now includes the **V** and **VCORR** matrices, which put **G** and **R** (or **GCORR** and **RCORR**, respectively) back together again. Thus, the structure we end up with in **V** after combining the random intercept variance and residual variance and covariance matrix is the <u>same compound symmetry as in an R-only model</u>.

<u>Random Intercept in G + Diagonal R Model</u> → Same as compound symmetry R-only model

```
TITLE1 "SAS Random Intercept + Diagonal R Model";
TITLE2 "Equal to Compound Symmetry R-only Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     RANDOM INTERCEPT / TYPE=UN V VCORR G SUBJECT=subid;
     REPEATED week / TYPE=VC R RCORR SUBJECT=subid;
RUN; TITLE1; TITLE2;
```

```
display "STATA Random Intercept + Diagonal R Model"
display "Equal to Compound Symmetry R-only Model"
mixed severity ,  || subid:  ,                          ///
         variance reml residuals(independent,t(week)) ///
         dfmethod(satterthwaite) dftable(pvalue)
      estat recovariance, relevel(subid) // G matrix
      estat wcorrelation, covariance     // V matrix
      estat wcorrelation                 // VCORR matrix
      estat ic, n(124)                   // AIC and BIC
      estat icc                          // ICC
```

In STATA, we remove the "noconstant" option in the random part to include a random intercept variance in the **G** matrix. Consequently, the "wcorrelation" option now refers to the **V** matrix instead of the **R** matrix.

```
print("R LME Random Intercept + Diagonal R Model")
RI = lme(data=Example4, method="REML", severity~1, random=~1|subid, correlation=NULL)
print("Show results using incorrect DDF"); summary(RI)
print("Show G, R, and V matrices for first person")
G=getVarCov(RI, individual="100", type="random.effects"); G
R=getVarCov(RI, individual="100", type="conditional"); R
V=getVarCov(RI, individual="100", type="marginal"); V
```

```
ICC=(V[[1]][2,1])/(V[[1]][1,1]); print("Show ICC"); ICC
print("LRT: Is RI worse than UN?"); anova(UN,RI)
```

SAS Output:

Estimated R Matrix for SUBID 100 → WP residual (conditional) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	0.1306						
2		0.1306					
3			0.1306				
4				0.1306			
5					0.1306		
6						0.1306	
7							0.1306

Estimated R Correlation Matrix for SUBID 100 → WP residual (conditional) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000						
2		1.0000					
3			1.0000				
4				1.0000			
5					1.0000		
6						1.0000	
7							1.0000

Estimated G Matrix

Row	Effect	subject id	Col1
1	Intercept	100	0.6821

Estimated V Matrix for SUBID 100 → TOTAL (marginal) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	0.8127	0.6821	0.6821	0.6821	0.6821	0.6821	0.6821
2	0.6821	0.8127	0.6821	0.6821	0.6821	0.6821	0.6821
3	0.6821	0.6821	0.8127	0.6821	0.6821	0.6821	0.6821
4	0.6821	0.6821	0.6821	0.8127	0.6821	0.6821	0.6821
5	0.6821	0.6821	0.6821	0.6821	0.8127	0.6821	0.6821
6	0.6821	0.6821	0.6821	0.6821	0.6821	0.8127	0.6821
7	0.6821	0.6821	0.6821	0.6821	0.6821	0.6821	0.8127

Estimated VCORR Matrix for SUBID 100 → TOTAL (marginal) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8393	0.8393	0.8393	0.8393	0.8393	0.8393
2	0.8393	1.0000	0.8393	0.8393	0.8393	0.8393	0.8393
3	0.8393	0.8393	1.0000	0.8393	0.8393	0.8393	0.8393
4	0.8393	0.8393	0.8393	1.0000	0.8393	0.8393	0.8393
5	0.8393	0.8393	0.8393	0.8393	1.0000	0.8393	0.8393
6	0.8393	0.8393	0.8393	0.8393	0.8393	1.0000	0.8393
7	0.8393	0.8393	0.8393	0.8393	0.8393	0.8393	1.0000

The intraclass correlation in VCORR can be computed as:

$$\text{ICC} = \frac{0.6820}{0.6820 + 0.1306} = .8393$$

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z	
UN(1,1)	SUBID	0.6821	0.09020	7.56	<.0001	L2 Random intercept U_0 variance
Week	SUBID	0.1306	0.007269	17.97	<.0001	L1 Residual e variance

```
                          Information Criteria
Neg2LogLike    Parms      AIC      AICC      HQIC       BIC      CAIC
     1049.7        2   1053.7    1053.7    1056.0    1059.4    1061.4

                 Solution for Fixed Effects
                            Standard
Effect        Estimate      Error      DF    t Value    Pr > |t|
Intercept       1.4767    0.07548     123      19.56      <.0001
```

We'll try **two types of residual correlations**: **AR1 and reduced Toeplitz** (covariances only at lags 5, 4, 3, 2, and 1).

<u>Random Intercept in G + AR1 R Model</u>: only residuals have AR1 correlation (equal residual variance)

```
TITLE1 "SAS Random Intercept + Auto-Regressive R Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     RANDOM INTERCEPT / TYPE=UN V VCORR G SUBJECT=subid;
     REPEATED week / TYPE=AR(1) R RCORR SUBJECT=subid;
     ODS OUTPUT InfoCrit=FitARRI; * Save for LRT;
RUN; TITLE1;
* Model comparison: did random intercept help?;
  %FitTest(FitFewer=FitAR, FitMore=FitARRI);
* Model comparison: not worse than UN?;
  %FitTest(FitFewer=FitARRI, FitMore=FitUN);
```

For unbalanced time, you can use `TYPE=SP(POW)(week)` instead

```
display "STATA Random Intercept + Auto-Regressive R Model"
mixed severity ,  || subid:  ,                     ///
         variance reml residuals(ar1,t(week))  ///
         dfmethod(satterthwaite) dftable(pvalue)
      estat recovariance, relevel(subid) // G matrix
      estat wcorrelation, covariance     // V matrix
      estat wcorrelation                 // VCORR matrix
      estat ic, n(124)                   // AIC and BIC
      estimates store FitARRI            // Save for LRT
      lrtest FitARRI FitAR               // Better than AR?
      lrtest FitUN FitARRI               // Worse than UN?
```

Option for unbalanced time are not available in STATA as far as I know

```
print("R LME Random Intercept + Auto-Regressive R Model")
ARRI = lme(data=Example4, method="REML", severity~1, random=~1|subid,
           correlation=(corAR1(form=~as.numeric(week1)|subid)))
print("Show results using incorrect DDF"); summary(ARRI)
print("Show G, R, RCORR, and V matrices for first person")
G=getVarCov(ARRI, individual="100", type="random.effects"); G
R=getVarCov(ARRI, individual="100", type="conditional"); R
RCORR=corMatrix(ARRI$modelStruct$corStruct)[[6]]; RCORR
V=getVarCov(ARRI, individual="100", type="marginal"); V
print("LRT: Is ARRI better than AR?"); anova(ARRI,AR)
print("LRT: Is ARRI worse than UN?"); anova(UN,ARRI)
```

For unbalanced time, you can use `corCAR1` instead

SAS Output:

Estimated R Matrix for SUBID 100 → WP residual (conditional) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1731**	0.08887	0.04564	0.02343	0.01203	0.006180	0.003173
2	0.08887	**0.1731**	0.08887	0.04564	0.02343	0.01203	0.006180
3	0.04564	0.08887	**0.1731**	0.08887	0.04564	0.02343	0.01203
4	0.02343	0.04564	0.08887	**0.1731**	0.08887	0.04564	0.02343
5	0.01203	0.02343	0.04564	0.08887	**0.1731**	0.08887	0.04564
6	0.006180	0.01203	0.02343	0.04564	0.08887	**0.1731**	0.08887
7	0.003173	0.006180	0.01203	0.02343	0.04564	0.08887	**0.1731**

Estimated R Correlation Matrix for SUBID 100 → WP residual (conditional) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.5135	0.2637	0.1354	0.06954	0.03571	0.01834
2	**0.5135**	1.0000	0.5135	0.2637	0.1354	0.06954	0.03571
3	**0.2637**	**0.5135**	1.0000	0.5135	0.2637	0.1354	0.06954
4	**0.1354**	**0.2637**	**0.5135**	1.0000	0.5135	0.2637	0.1354
5	**0.06954**	**0.1354**	**0.2637**	**0.5135**	1.0000	0.5135	0.2637
6	**0.03571**	**0.06954**	**0.1354**	**0.2637**	**0.5135**	1.0000	0.5135
7	**0.01834**	**0.03571**	**0.06954**	**0.1354**	**0.2637**	**0.5135**	1.0000

Estimated G Matrix → BP random intercept variance

Row	Effect	subid	Col1
1	Intercept	100	0.6350

Estimated V Matrix for SUBID 100 → TOTAL (marginal) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.8080**	0.7238	0.6806	0.6584	0.6470	0.6411	0.6381
2	0.7238	**0.8080**	0.7238	0.6806	0.6584	0.6470	0.6411
3	0.6806	0.7238	**0.8080**	0.7238	0.6806	0.6584	0.6470
4	0.6584	0.6806	0.7238	**0.8080**	0.7238	0.6806	0.6584
5	0.6470	0.6584	0.6806	0.7238	**0.8080**	0.7238	0.6806
6	0.6411	0.6470	0.6584	0.6806	0.7238	**0.8080**	0.7238
7	0.6381	0.6411	0.6470	0.6584	0.6806	0.7238	**0.8080**

Estimated V Correlation Matrix for SUBID 100 → TOTAL (marginal) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8958	0.8423	0.8148	0.8007	0.7935	0.7898
2	**0.8958**	1.0000	0.8958	0.8423	0.8148	0.8007	0.7935
3	**0.8423**	**0.8958**	1.0000	0.8958	0.8423	0.8148	0.8007
4	**0.8148**	**0.8423**	**0.8958**	1.0000	0.8958	0.8423	0.8148
5	**0.8007**	**0.8148**	**0.8423**	**0.8958**	1.0000	0.8958	0.8423
6	**0.7935**	**0.8007**	**0.8148**	**0.8423**	**0.8958**	1.0000	0.8958
7	**0.7898**	**0.7935**	**0.8007**	**0.8148**	**0.8423**	**0.8958**	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	SUBID	0.6350	0.09026	7.04	<.0001	L2 Random Intercept U_0 variance
AR(1)	SUBID	0.5135	0.05179	9.92	<.0001	L1 Auto-regressive correlation in R
Residual		0.1731	0.01777	9.74	<.0001	L1 Residual e variance

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
943.8	3	949.8	949.9	953.3	958.3	961.3

The AIC doesn't like this model better than our best model so far (**R**-only homogeneous Toeplitz), but the BIC could go either way.

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.4885	0.07515	122	19.81	<.0001

Adding the random intercept variance is a significant improvement over the **R**-only AR1, −2ΔLL(1) = 47.1, p < 001, indicating only the residuals need an AR correlation.

Likelihood Ratio Test for FitAR vs. FitARRI

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitAR	990.9	2	994.9	1000.5	.	.	.
FitARRI	943.8	3	949.8	958.3	47.0541	1	6.9053E-12

Likelihood Ratio Test for FitAR1RI vs. FitUN

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitAR1RI	943.8	3	949.8	958.3	.	.	.
FitUN	893.1	28	949.1	1028.1	50.6930	25	.001748946

AR+RI still fits worse than UN, −2ΔLL(25) = 50.7, p < 002.

<u>Random Intercept in G with Toeplitz(6) in R Model</u> (allows 5 lagged covariances; sets the last one to 0)

```
TITLE1 "SAS Random Intercept + 5-Lag Toeplitz6 R Model";
TITLE2 "Equal to Toeplitz (n=7 bands) R-only Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     RANDOM INTERCEPT / TYPE=UN V VCORR G SUBJECT=subid;
     REPEATED week / TYPE=TOEP(6) R RCORR SUBJECT=subid;
RUN; TITLE1; TITLE2;

display "STATA Random Intercept + 5-lag Toeplitz6 R Model"
display "Equal to Toeplitz (n=7 bands) R-only Model"
mixed severity ,  || subid:  ,                      ///
         variance reml residuals(toeplitz5,t(week)) ///
         dfmethod(satterthwaite) dftable(pvalue)
      estat recovariance, relevel(subid) // G matrix
      estat wcorrelation, covariance     // V matrix
      estat wcorrelation                 // VCORR matrix
      estat ic, n(124)                   // AIC and BIC

# Toeplitz is not a pre-defined structure in R lme
```

SAS Output:

Estimated R Matrix for SUBID 100 → WP residual (conditional) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1562**	0.07152	0.04321	0.003458	-0.01479	-0.01730	
2	0.07152	**0.1562**	0.07152	0.04321	0.003458	-0.01479	-0.01730
3	0.04321	0.07152	**0.1562**	0.07152	0.04321	0.003458	-0.01479
4	0.003458	0.04321	0.07152	**0.1562**	0.07152	0.04321	0.003458
5	-0.01479	0.003458	0.04321	0.07152	**0.1562**	0.07152	0.04321
6	-0.01730	-0.01479	0.003458	0.04321	0.07152	**0.1562**	0.07152
7		-0.01730	-0.01479	0.003458	0.04321	0.07152	**0.1562**

Estimated R Correlation Matrix for SUBID 100 → WP (conditional) residual correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.4579	0.2766	0.02214	-0.09467	-0.1108	
2	0.4579	1.0000	0.4579	0.2766	0.02214	-0.09467	-0.1108
3	0.2766	0.4579	1.0000	0.4579	0.2766	0.02214	-0.09467
4	0.02214	0.2766	0.4579	1.0000	0.4579	0.2766	0.02214
5	-0.09467	0.02214	0.2766	0.4579	1.0000	0.4579	0.2766
6	-0.1108	-0.09467	0.02214	0.2766	0.4579	1.0000	0.4579
7		-0.1108	-0.09467	0.02214	0.2766	0.4579	1.0000

Estimated G Matrix → BP random intercept variance

Row	Effect	subject id	Col1
1	Intercept	100	0.6541

Estimated V Matrix for SUBID 100 → TOTAL (marginal) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	0.8104	0.7257	0.6974	0.6576	0.6394	0.6368	0.6541
2	0.7257	0.8104	0.7257	0.6974	0.6576	0.6394	0.6368
3	0.6974	0.7257	0.8104	0.7257	0.6974	0.6576	0.6394
4	0.6576	0.6974	0.7257	0.8104	0.7257	0.6974	0.6576
5	0.6394	0.6576	0.6974	0.7257	0.8104	0.7257	0.6974
6	0.6368	0.6394	0.6576	0.6974	0.7257	0.8104	0.7257
7	0.6541	0.6368	0.6394	0.6576	0.6974	0.7257	0.8104

```
               Estimated V Correlation Matrix for SUBID 100 → TOTAL (marginal) correlation
 Row        Col1        Col2        Col3        Col4        Col5        Col6        Col7
   1      1.0000      0.8955      0.8606      0.8115      0.7890      0.7859      0.8072
   2      0.8955      1.0000      0.8955      0.8606      0.8115      0.7890      0.7859
   3      0.8606      0.8955      1.0000      0.8955      0.8606      0.8115      0.7890
   4      0.8115      0.8606      0.8955      1.0000      0.8955      0.8606      0.8115
   5      0.7890      0.8115      0.8606      0.8955      1.0000      0.8955      0.8606
   6      0.7859      0.7890      0.8115      0.8606      0.8955      1.0000      0.8955
   7      0.8072      0.7859      0.7890      0.8115      0.8606      0.8955      1.0000

                  Covariance Parameter Estimates
                                  Standard         Z
Cov Parm     Subject    Estimate       Error     Value        Pr Z
UN(1,1)      SUBID        0.6541     0.09179      7.13      <.0001  L2 Random intercept U0 variance
TOEP(2)      SUBID       0.07152     0.02189      3.27      0.0011  L1 Residual e covariance for t-1
TOEP(3)      SUBID       0.04321     0.02130      2.03      0.0425  L1 Residual e covariance for t-2
TOEP(4)      SUBID      0.003458     0.02096      0.16      0.8690  L1 Residual e covariance for t-3
TOEP(5)      SUBID      -0.01479     0.01860     -0.80      0.4266  L1 Residual e covariance for t-4
TOEP(6)      SUBID      -0.01730     0.01487     -1.16      0.2444  L1 Residual e covariance for t-5
Residual                  0.1562     0.02178      7.17      <.0001  L1 Residual e variance (equal across weeks)

                            Information Criteria
Neg2LogLike     Parms       AIC        AICC        HQIC         BIC        CAIC
      926.0         7     940.0       940.1       948.0       959.7       966.7

                   Solution for Fixed Effects
                          Standard
Effect        Estimate       Error       DF     t Value     Pr > |t|
Intercept       1.4884     0.07533      122       19.76       <.0001
```

This RI + TOEP(n−1) 5-lag model is equivalent to the Toeplitz(n) 6-lag model without the random intercept, as shown by the fit statistics. But this RI+5lag model has an interpretational advantage: rather than testing whether the total (marginal) lagged covariance (in **V**) is different than 0, it tests whether the residual (conditional) lagged covariance (in **R**) is different from 0 *after taking out individual differences in the intercept*. It looks like only some of the lagged covariances are significant, so we can probably simplify the model without hurting fit.

Summary of sequential models taking out higher-lag covariances:

TOEP(5) → 4 lags only:	REML −2LL = 927.2	AIC = 939.2	BIC = 956.2
TOEP(4) → 3 lags only:	REML −2LL = 927.3	AIC = 937.3	BIC = 951.4 ← Best so far
TOEP(3) → 2 lags only:	REML −2LL = 933.6	AIC = 941.6	BIC = 952.8

Random Intercept in G with Toeplitz(4) in R Model (allows 3 lagged covariances; sets the others to 0)

```
TITLE1 "SAS Random Intercept + 3-Lag Toeplitz4 R Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     RANDOM INTERCEPT / TYPE=UN V VCORR G SUBJECT=subid;
     REPEATED week / TYPE=TOEP(4) R RCORR SUBJECT=subid;
     ODS OUTPUT InfoCrit=FitToep4RI; * Save for LRT;
RUN; TITLE1;
* Model comparison: not worse than UN?;
  %FitTest(FitFewer=FitToep4RI, FitMore=FitUN);
```

```
display "STATA Random Intercept + 3-lag Toeplitz4 R Model"
mixed severity ,  || subid:  ,                          ///
         variance reml residuals(toeplitz3,t(week)) ///
         dfmethod(satterthwaite) dftable(pvalue)
```

```
estat recovariance, relevel(subid) // G matrix
estat wcorrelation, covariance     // V matrix
estat wcorrelation                 // VCORR matrix
estat ic, n(124)                   // AIC and BIC
estimates store FitToep4RI         // Save for LRT
lrtest FitUN FitToep4RI            // Worse than UN?
```

SAS Output:

Estimated R Matrix for SUBID 100 → WP residual (conditional) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1702**	0.08576	0.05791	0.01872			
2	0.08576	**0.1702**	0.08576	0.05791	0.01872		
3	0.05791	0.08576	**0.1702**	0.08576	0.05791	0.01872	
4	0.01872	0.05791	0.08576	**0.1702**	0.08576	0.05791	0.01872
5		0.01872	0.05791	0.08576	**0.1702**	0.08576	0.05791
6			0.01872	0.05791	0.08576	**0.1702**	0.08576
7				0.01872	0.05791	0.08576	**0.1702**

Estimated R Correlation Matrix for SUBID 100 → WP residual (conditional) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.5038	0.3402	0.1100			
2	0.5038	1.0000	0.5038	0.3402	0.1100		
3	0.3402	0.5038	1.0000	0.5038	0.3402	0.1100	
4	0.1100	0.3402	0.5038	1.0000	0.5038	0.3402	0.1100
5		0.1100	0.3402	0.5038	1.0000	0.5038	0.3402
6			0.1100	0.3402	0.5038	1.0000	0.5038
7				0.1100	0.3402	0.5038	1.0000

Estimated G Matrix → BP random intercept variance

Row	Effect	subject id	Col1
1	Intercept	100	0.6395

Estimated V Matrix for SUBID 100 → TOTAL (marginal) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	0.8097	0.7252	0.6974	0.6582	0.6395	0.6395	0.6395
2	0.7252	0.8097	0.7252	0.6974	0.6582	0.6395	0.6395
3	0.6974	0.7252	0.8097	0.7252	0.6974	0.6582	0.6395
4	0.6582	0.6974	0.7252	0.8097	0.7252	0.6974	0.6582
5	0.6395	0.6582	0.6974	0.7252	0.8097	0.7252	0.6974
6	0.6395	0.6395	0.6582	0.6974	0.7252	0.8097	0.7252
7	0.6395	0.6395	0.6395	0.6582	0.6974	0.7252	0.8097

Estimated V Correlation Matrix for SUBID 100 → TOTAL (marginal) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8957	0.8613	0.8129	0.7898	0.7898	0.7898
2	0.8957	1.0000	0.8957	0.8613	0.8129	0.7898	0.7898
3	0.8613	0.8957	1.0000	0.8957	0.8613	0.8129	0.7898
4	0.8129	0.8613	0.8957	1.0000	0.8957	0.8613	0.8129
5	0.7898	0.8129	0.8613	0.8957	1.0000	0.8957	0.8613
6	0.7898	0.7898	0.8129	0.8613	0.8957	1.0000	0.8957
7	0.7898	0.7898	0.7898	0.8129	0.8613	0.8957	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	SUBID	0.6395	0.09004	7.10	<.0001	L2 Random intercept U_0 variance
TOEP(2)	SUBID	0.08576	0.01267	6.77	<.0001	L1 Residual e covariance for t-1
TOEP(3)	SUBID	0.05791	0.009852	5.88	<.0001	L1 Residual e covariance for t-2
TOEP(4)	SUBID	0.01872	0.007469	2.51	0.0122	L1 Residual e covariance for t-3
Residual		0.1702	0.01401	12.15	<.0001	L1 Residual e variance (equal across weeks)

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
927.3	5	937.3	937.4	943.1	951.4	956.4

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.4899	0.07528	122	19.79	<.0001

Likelihood Ratio Test for FitToep4RI vs. FitUN

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitToep4RI	927.3	5	937.3	951.4	.	.	.
FitUN	893.1	28	949.1	1028.1	34.1944	23	0.062399

One last thing to check: do the residual variances in **R** need to be heterogeneous?

Random Intercept in G with Heterogeneous Toeplitz(3) in R Model
(allows 3 lagged covariances; sets the others to 0, adds unequal residual variances across weeks)

```
TITLE1 "SAS Random Intercept + 3-Lag Heterogeneous Toeplitz4 R Model";
PROC MIXED DATA=work.Example4 COVTEST NOCLPRINT IC METHOD=REML;
     CLASS subid week;
     MODEL severity = / SOLUTION DDFM=Satterthwaite;
     RANDOM INTERCEPT / TYPE=UN V VCORR G SUBJECT=subid;
     REPEATED week / TYPE=TOEPH(4) R RCORR SUBJECT=subid;
     ODS OUTPUT InfoCrit=FitToep4HRI; * Save for LRT;
RUN; TITLE1;
* Model comparison: did het variance help?;
  %FitTest(FitFewer=FitToep4RI, FitMore=FitToep4HRI);
* Model comparison: now not worse than UN?;
  %FitTest(FitFewer=FitToep4HRI, FitMore=FitUN);
```

This model is only available directly in SAS MIXED as far as I can tell.

SAS Output:

Estimated R Matrix for SUBID 100 → WP (conditional) residual variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1261**	0.06946	0.06022	0.01485			
2	0.06946	**0.1462**	0.09301	0.05131	0.01985		
3	0.06022	0.09301	**0.2261**	0.09152	0.07918	0.02044	
4	0.01485	0.05131	0.09152	**0.1416**	0.08986	0.05187	0.01808
5		0.01985	0.07918	0.08986	**0.2180**	0.09231	0.07197
6			0.02044	0.05187	0.09231	**0.1494**	0.08546
7				0.01808	0.07197	0.08546	**0.1868**

Estimated R Correlation Matrix for SUBID 100 → WP (conditional) residual correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.5115	0.3566	0.1112			
2	0.5115	1.0000	0.5115	0.3566	0.1112		
3	0.3566	0.5115	1.0000	0.5115	0.3566	0.1112	
4	0.1112	0.3566	0.5115	1.0000	0.5115	0.3566	0.1112
5		0.1112	0.3566	0.5115	1.0000	0.5115	0.3566
6			0.1112	0.3566	0.5115	1.0000	0.5115
7				0.1112	0.3566	0.5115	1.0000

Residual *covariances are unequal* within a time lag band (because the residual variances are allowed to differ over weeks), but the residual *correlations are equal* within a band.

Estimated G Matrix → BP random intercept variance

Row	Effect	subject id	Col1
1	Intercept	100	0.6503

Estimated V Matrix for SUBID 100 → TOTAL (marginal) variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.7764**	0.7198	0.7105	0.6652	0.6503	0.6503	0.6503
2	0.7198	**0.7965**	0.7433	0.7016	0.6701	0.6503	0.6503
3	0.7105	0.7433	**0.8764**	0.7418	0.7295	0.6707	0.6503
4	0.6652	0.7016	0.7418	**0.7919**	0.7402	0.7022	0.6684
5	0.6503	0.6701	0.7295	0.7402	**0.8683**	0.7426	0.7223
7	0.6503	0.6503	0.6503	0.6684	0.7223	0.7358	**0.8371**

Estimated V Correlation Matrix for SUBID 100 → TOTAL (marginal) correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.9153	0.8614	0.8483	0.7920	0.8253	0.8066
2	0.9153	1.0000	0.8896	0.8834	0.8058	0.8148	0.7964
3	0.8614	0.8896	1.0000	0.8905	0.8362	0.8012	0.7592
4	0.8483	0.8834	0.8905	1.0000	0.8926	0.8824	0.8209
5	0.7920	0.8058	0.8362	0.8926	1.0000	0.8912	0.8472
6	0.8253	0.8148	0.8012	0.8824	0.8912	1.0000	0.8992
7	0.8066	0.7964	0.7592	0.8209	0.8472	0.8992	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	SUBID	0.6503	0.09069	7.17	<.0001	L2 Random intercept U_0 variance
Var(1)	SUBID	0.1261	0.02255	5.59	<.0001	L1 Residual e Variance at week 1
Var(2)	SUBID	0.1462	0.02283	6.41	<.0001	L1 Residual e Variance at week 2
Var(3)	SUBID	0.2261	0.03247	6.96	<.0001	L1 Residual e Variance at week 3
Var(4)	SUBID	0.1416	0.02442	5.80	<.0001	L1 Residual e Variance at week 4
Var(5)	SUBID	0.2180	0.03228	6.75	<.0001	L1 Residual e Variance at week 5
Var(6)	SUBID	0.1494	0.02411	6.20	<.0001	L1 Residual e Variance at week 6
Var(7)	SUBID	0.1868	0.02890	6.47	<.0001	L1 Residual e Variance at week 7
TOEPH(1)	SUBID	0.5115	0.03899	13.12	<.0001	L1 Residual e correlation for t-1
TOEPH(2)	SUBID	0.3566	0.03764	9.48	<.0001	L1 Residual e correlation for t-2
TOEPH(3)	SUBID	0.1112	0.04115	2.70	0.0069	L1 Residual 3 correlation for t-3

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
910.0	11	932.0	932.3	944.6	963.0	974.0

Likelihood Ratio Test for FitToep3RI vs. FitToep3HRI

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitToep4RI	927.3	5	937.3	951.4	.	.	.
FitToep4HRI	910.0	11	932.0	963.0	17.3234	6	.008165240

Likelihood Ratio Test for FitToep3HRI vs. FitUN

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitToep3HRI	910.0	11	932.0	963.0	.	.	.
FitUN	893.1	28	949.1	1028.1	16.8710	17	0.46314

Adding separate residual variances improves fit, $-2\Delta LL(6) = 17.3$, $p = .008$, although BIC disagrees.

The model now fits not worse than UN, $-2\Delta LL(17) = 16.8$, $p = .463$. We win! But this fit should be re-checked after adding predictors.

A results section for these analyses (that are not of substantive interest) would be just a few sentences:

To ensure accurate standard errors for all fixed effects, the adequacy of fit of the variance–covariance model across all occasions was first evaluated using likelihood ratio tests (i.e., the χ^2 for the −2LL difference between nested models given degrees of freedom equal to the number of additional model parameters). To maximize power, we sought to reduce the number of estimated parameters by examining alternative models. Fortunately, we found a more parsimonious model that showed no decrease in fit, $-2\Delta LL(17) = 16.8$, $p = .463$, which included a random intercept variance, separate residual variances per occasion, and Toeplitz residual correlations for outcomes 1, 2, and 3 weeks apart.